MAR 26 1921

Journal of Educational Research

Official Organ of the National Association of Directors of Educational Research

B. R. BUCKINGHAM, Editor
Associate Editors
E. J. Ashbaugh — W. W. Charters — George D. Strayer
Leonard P. Ayres — S. A. Courtis — Lewis M. Terman
Walter S. Monroe

VOL. III FEBRUARY, 1921 No. 2

CONTENTS

STANDARDIZED VOCABULARY AND SENTENCE TESTS IN FRENCH ..... V. A. C. Henmon 81

THE OTIS GROUP INTELLIGENCE SCALE APPLIED TO THE ELEMENTARY-SCHOOL GRADUATING CLASSES OF OAKLAND, CALIFORNIA .. V. E. Dickson and John K. Norton 106

THE VALUE OF THE WRITTEN PARAGRAPH SUMMARY ... Charles E. Germane 116

AGE AND GRADE NORMS FOR THE NATIONAL INTELLIGENCE TESTS, SCALES A AND B ..... L. M. Terman and Ethel D. Whitmire 124

THE VALUE OF A HANDWRITING SCALE TO AN UNTRAINED TEACHER ..... R. L. Morton 133

EDITORIALS 138
REVIEWS AND ABSTRACTS 144
NEWS ITEMS AND COMMUNICATIONS 147
NATIONAL ASSOCIATION OF DIRECTORS OF EDUCATIONAL RESEARCH 158

Entered as second-class matter January 13, 1920, at the Post Office at Menasha, Wisconsin, under the act of March 3, 1879.

AMERICAN BOOK COMPANY

EDUCATIONAL PUBLISHERS

NEW YORK, 100 Washington Square
CINCINNATI, 300 Pike Street
CHICAGO, 330 East 22nd Street
BOSTON, 63 Summer Street
ATLANTA, 2 North Forsyth Street

Address all communications to the nearest office of the company. Descriptive circulars, catalogues, and price lists will be mailed on request to any address.

Your correspondence is solicited and will have prompt attention.

THE ILLINOIS EXAMINATION

A' la carte!!

Those who desire the various parts of the now famous ILLINOIS EXAMINATION supplied separately will be glad to learn that we are now able supply the several parts separately as desired. Order the parts as follows:

Illinois General Intelligence Scale.....$2.00 per 100
Monroe General Survey Scale in Arithmetic..... 1.00 per 100
Monroe Standardized Silent Reading—Revised..... .80 per 100

Those who have used Form I of the ILLINOIS EXAMINATION will be glad to learn that Form II is now available in whole or in part as desired.

Orders filled promptly

By

THE PUBLIC SCHOOL PUBLISHING CO.

Bloomington, Illinois

JOURNAL *of* EDUCATIONAL RESEARCH

VOLUME III FEBRUARY, 1921 NUMBER 2

STANDARDIZED VOCABULARY AND SENTENCE TESTS IN FRENCH

V. A. C. HENMON

University of Wisconsin

The ultimate objectives of modern foreign language study are the ability to understand the language, written or spoken, and the ability to speak or write the language. The most obvious indices of ability to understand a foreign language would be found in the number of words whose meanings are known or whose equivalents in English can be given, and in the rate of accuracy of translating connected discourse from the mother tongue to the foreign language. Ability or progress in a foreign language may then be measured by determining (1) the scope and accuracy of vocabulary, and (2) the difficulty of sentences that are understood or can be translated into English. The present paper will report tests for measuring ability in French in these two directions. The method and technic are the same as used in developing the Henmon Vocabulary and Sentence Tests in Latin.[1] The method of scaling words or sentences in difficulty is given in full in these papers and has been reported in many other studies in recent years so often that the statistical discussion in this report will be much abbreviated. I am greatly indebted to my wife and to Miss Alice Jackson for the laborious task of tabulating and scoring the papers and to many principals and teachers of French for giving the preliminary tests.

VOCABULARY TESTS

The essential difficulty in the construction of vocabulary tests lies in the selection of words to be used. If results are to be strictly comparable beyond the limits of a single class or of classes using the same text, they must necessarily be such as the pupil

[1] Henmon, V. A. C. "The measurement of ability in Latin," *Journal of Educational Psychology*, 8:515–38, 589–99; 11:121–36, November-December, 1917; March, 1920.

has certainly come in contact with and might reasonably be expected to know. In order to overcome this difficulty and to secure a standard vocabulary a tabulation was made of all the words occurring in twelve recent or widely used first-year texts, selected by competent teachers of French. The result was a list of 448 words out of approximately 4,000 words common to all the books. While these 448 words will occur with varying frequency in the different texts, the chances are slight that, if any other book is used than one of the twelve from which the tabulations were made, these words would not be found. The list, therefore, forms in a real sense a standard vocabulary.

The 448 words were divided into four lists of 112 words each. These lists appear in Table II. They were given to first-, second-, and third-year pupils in sixteen schools distributed in six states. They were given also to pupils who had had one-half, one and a half, and two and a half years of French, but the number of schools is too small for use in determining scale values. Similarly the number of schools and of pupils in the third year is too small for satisfactory use. Table I gives the results, therefore, for first- and second-year pupils only, both the averages and medians being given for the convenience of those who may wish to give the original lists rather than the tests derived from them. The difference between the two measures of central tendency is sometimes fairly large where the number of cases is small or the skewness of the distribution is marked. In scoring the words in the preliminary tests any translation given in a standard dictionary was accepted. Where words such as "lit," "la," "vers," and "montre" had several meanings, any one was scored as correct.

The averages and medians in Table I were obtained by scoring the number of words correct regardless of the variation in the difficulty of the words. It is, of course, inaccurate to assign the same value to *animal*, which is never missed by either first or second-year pupils, as to *afin que*, which is missed by 78 percent of first-year, and 76.8 percent of second-year pupils. The purpose of the study was to determine the weight or values to be assigned to each word for each year and a general scale value or weight. The information necessary for this purpose appears in Table II which gives the percent of times each word was correctly translated in the first and second years. These percents cannot be used directly since it is necessary to make allowance for the fact

TABLE I. RESULTS IN THE VOCABULARY TESTS, ALL SCHOOLS COMBINED

Lists	First-Year Pupils			Second-Year Pupils		
	No. of Pupils	Averages	Medians	No. of Pupils	Averages	Medians
I	415	82.0	85	95	100.0	102
II	517	76.5	79	70	95.3	98
III	440	76.8	78	109	94.9	99
IV	482	74.2	77	125	93.8	96
All.....	1,854	77.3		399	95.0	

that the distribution of mental abilities is normal, or in accordance with the bellshaped curve of distribution, and not rectangular. The method by which this is done consists in substance of locating each word on the base line of a normal surface of frequency in terms of the probable error equivalents of the percent of times it is correctly translated. In order to eliminate minus values an approximation to a zero point of ability was arbitrarily taken by locating the position of a word so easy as to be missed by but five out of one thousand first-year pupils. This point lies 3.8 times the probable error below the median. A word so difficult as to be translated correctly by but five out of one thousand pupils would be located at +3.8 P. E. The scale values which appear in Table II were determined by subtracting from or adding to 3.8 P. E., the P. E. equivalents of the percent of times each word was correctly translated.

The derivation of the general scale value required the determination of the size of the intervals between the medians of the first and second years, and this requires the amount and percent of overlapping between the two years.

The details of these calculations are omitted. It will suffice to state that the interval between the first and second years expressed in terms of the probable error as the unit of measurement is 2.315 P. E. The average position of the zero point is therefore .−2.642 P. E. which becomes the point of reference to which each word is referred. The general scale values thus derived represent the best approximation to an average weight to be assigned regardless of year or grade.

TABLE II. TOTAL LIST OF 448 WORDS GIVEN PUPILS

List 1	Percents Correct		Year Scale Values		General Scale Values
	I	II	I	II	
1. à	96.1	98.9	1.2	0.4	0.8
2. âge	88.7	100.0	2.0	0.0	1.0
3. alors	81.9	96.8	2.4	1.1	1.7
4. ancien	94.0	97.9	1.5	0.8	1.1
5. appartenir	11.3	18.9	5.6	5.1	5.3
6. après	72.5	86.3	2.9	2.2	2.5
7. arriver	98.3	98.9	0.7	0.4	0.5
8. aussi	99.0	96.8	0.3	1.1	0.7
9. autre	90.6	96.8	1.8	1.1	1.4
10. avoir	90.1	98.9	1.9	0.4	1.1
11. bien	97.4	100.0	0.9	0.0	0.4
12. bleu	95.7	98.9	1.3	0.4	0.8
13. bout	32.8	60.0	4.5	3.4	3.9
14. café	96.6	97.9	1.1	0.8	0.9
15. ceci	62.7	75.8	3.3	2.8	3.0
16. cependant	34.7	83.1	4.4	2.4	3.4
17. champ	70.4	91.5	3.0	1.8	2.4
18. chaud	59.8	85.2	3.4	2.3	2.8
19. cheval	91.3	98.9	1.8	0.4	1.1
20. chose	77.1	97.9	2.7	0.8	1.7
21. classe	97.6	100.0	0.9	0.0	0.4
22. commencer	98.1	98.9	0.7	0.4	0.5
23. connaître	54.7	96.8	3.6	1.1	2.3
24. côté	51.6	90.5	3.7	1.9	2.8
25. courir	54.2	91.5	3.6	1.8	2.7
26. craindre	22.4	57.9	4.9	3.5	4.2
27. dans	97.4	100.0	0.9	0.0	0.4
28. déjà	68.7	87.3	3.1	2.1	2.6
29. demi	74.9	97.9	2.8	0.8	1.8
30. descendre	95.2	100.0	1.3	0.0	0.6
31. devant	64.8	90.5	3.2	1.9	2.5
32. dîner	91.1	91.5	1.7	1.8	1.7
33. droit (adj.)	59.3	92.6	3.5	1.7	2.6
34. élève	84.3	84.2	2.3	2.3	2.3
35. encore	93.5	95.8	1.6	1.3	1.4
36. entre	43.8	80.0	4.0	2.6	3.3
37. et	99.0	100.0	0.3	0.0	0.1
38. faire	92.9	100.0	1.6	0.0	0.8

TABLE II—CONTINUED

List 1—Continued	Percents Correct		Year Scale Values		General Scale Values
	I	II	I	II	
39. fenêtre	97.2	97.9	1.0	0.8	0.9
40. fils	73.0	96.8	2.9	1.1	2.0
41. fois	67.4	97.9	3.1	0.8	1.9
42. franc (n.)	86.9	95.8	2.1	1.3	1.7
43. fruit	93.3	97.9	1.6	0.8	1.2
44. gens	53.3	85.2	3.7	2.3	3.0
45. guerre	75.2	93.7	2.8	1.5	2.1
46. heureux	82.4	98.9	2.4	0.4	1.4
47. homme	99.0	100.0	0.3	0.0	0.1
48. inviter	96.1	100.0	1.2	0.0	0.6
49. jeter	51.6	89.4	3.7	1.9	2.8
50. journal	89.2	95.8	2.0	1.3	1.6
51. laisser	60.0	83.1	3.4	2.4	2.9
52. le	95.9	100.0	1.2	0.0	0.6
53. leur	89.2	97.9	2.0	0.8	1.4
54. lire	51.6	96.8	3.7	1.1	2.4
55. loin	18.9	61.1	5.1	3.4	4.2
56. madame	90.3	98.9	1.9	0.4	1.1
57. mais	99.2	100.0	0.2	0.0	0.1
58. manger	93.5	97.9	1.6	0.8	1.2
59. me	88.7	91.5	2.0	1.8	1.9
60. mer	65.3	83.1	3.2	2.4	2.8
61. mieux	55.9	84.2	3.6	2.3	2.9
62. moins	56.1	87.4	3.6	2.1	2.8
63. monde	88.2	97.9	2.0	0.8	1.4
64. montrer	68.7	66.3	3.1	3.2	3.1
65. mourir	58.8	90.5	3.5	1.9	2.7
66. ni	70.8	89.4	3.0	1.9	2.4
67. nôtre (pron.)	82.9	94.7	2.4	1.4	1.9
68. nuit	81.7	97.9	2.5	0.8	1.6
69. offrir	88.2	100.0	2.0	0.0	1.0
70. oreille	33.1	68.4	4.4	3.1	3.7
71. ou	63.9	64.2	3.3	3.3	3.3
72. ouvrir	73.3	91.5	3.0	1.8	2.4
73. paraître	12.0	70.5	5.5	3.0	4.2
74. partie	48.7	74.7	3.8	2.8	3.3
75. passer	92.7	96.8	1.6	1.1	1.3

TABLE II—CONTINUED

List 1—Continued	Percents Correct		Year Scale Values		General Scale Values
	I	II	I	II	
76. pendant	48.9	83.1	3.8	2.4	3.1
77. permettre	69.1	94.7	3.1	1.4	2.2
78. peu	85.5	98.9	2.2	0.4	1.3
79. plein	35.9	80.0	4.3	2.6	3.4
80. plusieurs	53.3	73.7	3.7	2.9	3.3
81. porter	82.9	94.7	2.4	1.4	1.9
82. pouvoir	58.3	91.5	3.5	1.8	2.6
83. presque	35.2	76.8	4.4	2.7	3.5
84. promener	70.6	90.5	3.0	1.9	2.4
85. quatre	96.8	97.9	1.1	0.8	0.9
86. que (pron.)	89.4	90.5	1.9	1.9	1.9
87. qui	92.8	97.9	1.6	0.8	1.2
88. quoique	14.9	25.3	5.3	4.8	5.0
89. rentrer	47.9	62.1	3.9	3.3	3.6
90. rendre	60.2	70.5	3.4	3.0	3.2
91. robe	93.0	97.9	1.6	0.8	1.2
92. sans	79.5	96.8	2.6	1.1	1.8
93. second	89.9	97.9	1.9	0.8	1.3
94. sept	93.5	98.9	1.6	0.4	1.0
95. si	92.3	98.9	1.7	0.4	1.0
96. soleil	88.9	95.8	2.0	1.3	1.6
97. sou	62.4	76.8	3.3	2.7	3.0
98. souvent	53.1	82.1	3.7	2.4	3.0
99. suivre	24.8	87.4	4.8	2.1	3.4
100. tant	26.3	68.4	4.7	3.1	3.9
101. temps	94.2	95.8	1.5	1.3	1.4
102. tomber	77.7	97.9	2.7	0.8	1.7
103. tout (adv.)	91.3	85.2	1.8	2.3	2.0
104. très	96.1	98.9	1.2	0.4	0.8
105. trouver	88.5	100.0	2.0	0.0	1.0
106. vache	80.9	82.1	2.6	2.4	2.5
107. vent	62.4	78.9	3.3	2.6	2.9
108. viande	75.2	86.3	2.8	2.2	2.5
109. vingt	88.7	94.7	2.0	1.4	1.7
110. vivre	74.7	93.7	2.8	1.5	2.1
111. voiture	22.3	93.7	4.9	1.5	3.2
112. vous	97.2	95.8	1.0	1.3	1.1

TABLE II—CONTINUED

List 2	Percents Correct		Year Scales Values		General Scale Values
	I	II	I	II	
1. acheter	95.0	97.1	1.4	1.0	1.2
2. aimer	98.3	100.0	0.7	0.0	0.3
3. ami	98.3	100.0	0.7	0.0	0.3
4. animal	100.0	100.0	0.0	0.0	0.0
5. appeler	79.3	100.0	2.6	0.0	1.3
6. arbre	97.9	97.1	0.8	1.0	0.9
7. assez	59.6	80.0	3.4	2.6	3.0
8. aussitôt	28.0	37.1	4.7	4.3	4.5
9. avancer	83.6	92.9	2.4	1.6	2.0
10. beau	98.6	100.0	0.5	0.0	0.2
11. bientôt	49.9	52.9	3.8	3.7	3.7
12. boire	43.7	74.3	4.0	2.8	3.4
13. bras	55.5	91.4	3.6	1.8	2.7
14. campagne	88.6	91.4	2.0	1.8	1.9
15. cela	72.1	87.1	2.9	2.1	2.5
16. chacun	36.3	88.6	4.3	2.0	3.1
17. chanter	91.3	97.1	1.8	1.0	1.4
18. chemin	54.9	91.4	3.6	1.8	2.7
19. chez	53.6	47.1	3.7	3.9	3.8
20. ciel	58.0	92.9	3.5	1.6	2.5
21. cœur	72.5	100.0	2.9	0.0	1.4
22. comment	64.4	94.3	3.3	1.5	2.4
23. content	93.0	100.0	1.6	0.0	0.8
24. coucher	61.1	91.4	3.4	1.8	2.6
25. court	23.8	45.7	4.9	4.0	4.4
26. crier	81.8	95.7	2.5	1.3	1.9
27. de	94.8	98.6	1.4	0.5	0.9
28. déjeuner (n.)	32.3	45.7	4.5	4.0	4.2
29. depuis	29.2	77.1	4.6	2.7	3.6
30. désirer	99.6	100.0	0.0	0.0	0.0
31. devenir	20.3	47.1	5.0	3.9	4.4
32. dont	36.9	55.7	4.3	3.6	3.8
33. eau	83.4	97.1	2.4	1.0	1.7
34. elle	98.6	100.0	0.5	0.0	0.2
35. enfant	96.3	100.0	1.2	0.0	0.6
36. entrer	96.5	98.6	1.1	0.5	0.8
37. être	91.7	98.6	1.7	0.5	1.1
38. falloir	21.3	52.9	5.0	3.7	4.3

TABLE II—CONTINUED

List 2—Continued	Percents Correct		Year Scale Values		General Scale Values
	I	II	I	II	
39. fer	42.9	67.1	4.1	3.1	3.6
40. fin (n.)	47.0	85.7	3.9	2.2	3.0
41. fort	74.7	94.3	2.8	1.5	2.1
42. frapper	77.0	95.7	2.7	1.3	2.0
43. gagner	35.2	85.7	4.4	2.2	3.3
44. grammaire	97.7	97.1	0.8	1.0	0.9
45. habit	42.9	67.1	4.1	3.1	3.6
46. hier	74.3	91.4	2.8	1.8	2.3
47. huit	75.6	97.1	2.8	1.0	1.9
48. jamais	82.4	100.0	2.4	0.0	1.1
49. jeune	91.5	97.1	1.8	1.0	1.4
50. journée	22.1	65.7	4.9	3.2	4.0
51. lait	68.3	88.6	3.1	2.0	2.5
52. leçon	94.2	97.1	1.5	1.0	1.2
53. lever	59.8	97.1	3.4	1.0	2.2
54. lit	38.3	85.7	4.2	2.2	3.2
55. long	87.0	98.6	2.1	0.5	1.3
56. magasin	84.9	92.9	2.3	1.6	1.9
57. maison	99.4	97.1	0.1	1.0	0.5
58. marcher	88.4	100.0	2.0	0.0	1.0
59. meilleur	69.6	91.4	3.0	1.8	2.4
60. mère	98.8	100.0	0.4	0.0	0.2
61. mille	71.8	95.7	2.9	1.3	2.1
62. moindre	25.5	37.1	4.8	4.3	4.5
63. monsieur	98.1	100.0	0.7	0.0	0.3
64. morceau	77.2	91.4	2.7	1.8	2.2
65. ne	76.6	87.1	2.7	2.1	2.4
66. noir	76.0	85.7	2.8	2.2	2.5
67. notre (adj.)	78.9	91.4	2.6	1.8	2.2
68. nul	26.3	31.4	4.7	4.5	4.6
69. on	67.1	82.9	3.1	2.4	2.7
70. ordonner	44.9	68.6	4.0	3.1	3.5
71. oublier	43.5	90.0	4.0	1.9	2.9
72. pain	91.5	98.6	1.8	0.5	1.1
73. parce que	77.8	98.6	2.7	0.5	1.6
74. partir	66.1	95.7	3.2	1.3	2.2
75. pauvre	93.2	100.0	1.6	0.0	0.8

TABLE II—CONTINUED

List 2—Continued	Percents Correct		Year Scales Values		General Scale Values
	I	II	I	II	
76. penser	80.7	100.0	2.5	0.0	1.2
77. personne (n.)	57.5	52.9	3.5	3.7	3.6
78. peur	27.0	80.0	4.7	2.6	3.6
79. pleurer	32.7	70.0	4.5	3.0	3.7
80. point (adv.)	10.6	34.3	5.7	4.4	5.0
81. pour	89.4	91.4	1.9	1.8	1.8
82. premier	95.6	98.6	1.3	0.5	0.9
83. prêt	41.0	62.9	4.1	3.3	3.7
84. puis	51.8	70.0	3.8	3.0	3.4
85. que (conj.)	76.4	88.6	2.7	2.0	2.3
86. quinze	75.8	87.1	2.8	2.1	2.4
87. recevoir	89.8	98.6	1.9	0.5	1.2
88. remarquer	88.4	91.4	2.0	1.8	1.9
89. répondre	92.6	100.0	1.7	0.0	0.8
90. riche	98.3	100.0	0.7	0.0	0.3
91. rouge	97.7	98.6	0.8	0.5	0.6
92. sauver	38.3	62.9	4.2	3.3	3.7
93. semaine	68.7	92.9	3.1	1.6	2.3
94. servir	77.0	97.1	3.7	1.0	2.3
95. six	97.3	100.0	0.9	0.0	0.4
96. soir	68.5	92.9	3.1	1.6	2.3
97. soulier	26.7	62.9	4.7	3.3	4.0
98. sucre	71.0	88.6	3.0	2.0	2.5
99. sur	90.9	80.0	1.8	2.6	2.2
100. tante	61.7	82.9	3.4	2.4	2.9
101. tenir	50.3	85.7	3.8	2.2	3.0
102. ton	45.3	81.4	4.0	2.5	3.2
103. tout (adj.)	94.6	95.7	1.4	1.3	1.3
104. trois	96.1	98.6	1.2	0.5	0.8
105. tu	82.8	95.7	2.4	1.3	1.8
106. valoir	9.0	20.0	5.8	5.0	5.4
107. verre	36.6	67.1	4.3	3.1	3.7
108. vie	65.6	92.9	3.2	1.6	2.4
109. vieux	72.5	95.7	2.9	1.3	2.1
110. voici	68.9	78.6	3.1	2.6	2.8
111. voix	57.8	88.6	3.5	2.0	2.7
112. voyage	96.7	98.6	1.1	0.5	0.8

TABLE II—CONTINUED

List 3	Percents Correct		Year Scale Values		General Scale Values
	I	II	I	II	
1. affaire	88.6	90.8	2.0	1.8	1.9
2. ainsi	43.0	66.1	4.1	3.2	3.6
3. amuser	98.4	94.7	0.6	1.4	1.0
4. année	83.2	84.4	2.4	2.3	2.3
5. apporter	81.6	85.3	2.5	2.3	2.4
6. argent	95.7	99.1	1.3	0.3	0.8
7. attendre	55.0	84.4	3.6	2.3	2.9
8. autant	19.5	30.3	5.1	4.6	4.8
9. avant	62.1	89.9	3.3	1.9	2.6
10. beaucoup	93.6	95.4	1.5	1.3	1.4
11. billet	54.1	92.7	3.6	1.6	2.6
12. bois	67.9	95.4	3.1	1.3	2.2
13. bruit	53.4	80.7	3.7	2.5	3.1
14. car	55.9	83.5	3.6	2.4	3.0
15. celui	48.2	45.9	3.9	4.0	3.9
16. chambre	93.6	99.1	1.5	0.3	0.9
17. chapeau	93.4	98.2	1.6	0.7	1.1
18. cher	92.9	99.1	1.6	0.3	0.9
19. chien	88.2	94.2	2.0	1.5	1.7
20. cinquante	57.5	77.1	3.5	2.7	3.1
21. combien	83.4	85.3	2.4	2.3	2.3
22. comprendre	77.0	94.5	2.7	1.4	2.0
23. continuer	96.6	99.1	1.1	0.3	0.7
24. coup	25.4	61.5	4.8	3.4	4.1
25. couteau	37.5	45.9	4.3	4.0	4.1
26. croire	22.9	85.3	4.9	2.3	3.6
27. décider	95.5	100.0	1.3	0.0	0.6
28. demain	67.0	93.6	3.1	1.5	2.3
29. dernier	47.9	91.6	3.9	1.8	2.8
30. deux	98.2	99.1	0.7	0.3	0.5
31. devoir	67.4	68.8	3.1	3.1	3.1
32. dormir	59.3	94.5	3.5	1.4	2.4
33. école	97.0	95.4	1.0	1.3	1.1
34. en (adv.)	37.7	43.1	4.3	4.1	4.2
35. enfin	46.9	83.5	3.9	2.4	3.1
36. envoyer	39.1	75.2	4.2	2.8	3.5
37. étudier	94.3	99.1	1.5	0.3	0.9
38. famille	88.2	95.4	2.0	1.3	1.6

TABLE II—CONTINUED

List 3—Continued	Percents Correct		Year Scale Values		General Scale Values
	I	II	I	II	
39. feu	52.7	90.8	3.7	1.8	2.7
40. finir	94.8	97.2	1.4	1.0	1.2
41. fou, fol	28.9	71.6	4.6	3.0	3.8
42. frère	95.7	99.1	1.3	0.3	0.8
43. garçon	98.8	100.0	0.4	0.0	0.2
44. grand	97.8	100.0	0.9	0.0	0.4
45. haut	70.0	91.7	3.0	1.7	2.3
46. hiver	86.6	91.7	2.2	1.7	1.9
47. ici	95.9	96.3	1.2	1.2	1.2
48. jardin	98.4	96.3	0.6	1.2	0.9
49. jouer	80.0	88.1	2.6	2.1	2.3
50. la	81.6	81.6	2.5	2.5	2.5
51. langue	86.6	92.7	2.2	1.6	1.9
52. lequel	54.5	74.3	3.6	2.8	3.2
53. lieu	6.8	32.1	6.0	4.5	5.2
54. livre	95.2	98.2	1.3	0.7	1.0
55. lorsque	31.8	54.1	4.5	3.6	4.0
56. main	85.9	100.0	2.2	0.0	1.1
57. maître	47.3	84.4	3.9	2.3	3.1
58. matin	95.0	99.1	1.4	0.3	0.8
59. même	80.9	93.5	2.5	1.6	2.0
60. mettre	36.8	87.2	4.3	2.1	3.2
61. minuit	13.6	45.0	5.4	4.0	4.7
62. mois	53.5	86.2	3.7	2.2	2.9
63. monter	81.1	62.4	2.5	3.3	2.9
64. mort	80.0	89.9	2.6	1.9	2.2
65. neige	82.4	89.0	2.4	2.0	2.2
66. nom	86.1	98.2	2.2	0.7	1.4
67. nous	91.1	94.5	1.8	1.4	1.6
68. obéir	72.7	91.7	2.9	1.7	2.3
69. onze	81.8	93.5	2.5	1.6	2.0
70. ordre	51.4	47.7	3.7	3.9	3.8
71. oui	97.0	98.2	1.0	0.7	0.8
72. papier	99.0	100.0	0.3	0.0	0.1
73. parler	98.6	100.0	0.5	0.0	0.2
74. pas (adv.)	82.0	93.6	2.4	1.5	1.9
75. payer	86.4	95.4	2.2	1.3	1.7

TABLE II—CONTINUED

List 3—Continued	Percents Correct		Year Scale Values		General Scale Values
	I	II	I	II	
76. père	99.0	98.2	0.3	0.7	0.5
77. personne (pron.)	48.2	71.6	3.9	3.0	3.4
78. pied	74.1	85.3	2.8	2.3	2.5
79. plume	95.0	100.0	1.4	0.0	0.7
80. point (n.)	47.5	68.8	3.9	3.1	3.5
81. pourquoi	92.5	96.3	1.7	1.2	1.4
82. prendre	47.3	90.8	3.9	1.8	2.8
83. printemps	66.8	72.5	3.2	2.9	3.0
84. quand	95.5	99.1	1.3	0.3	0.8
85. quelque	61.1	67.9	3.4	3.1	3.2
86. quitter	57.0	86.2	3.0	2.2	2.6
87. regarder	92.9	99.1	1.6	0.3	0.9
88. remercier	71.1	89.9	3.0	1.9	2.4
89. rester	52.9	82.6	3.7	2.4	3.0
90. rien	68.2	89.9	3.1	1.9	2.5
91. rue	87.0	94.5	2.1	1.4	1.7
92. savoir	57.9	94.5	3.5	1.4	2.4
93. sembler	40.9	85.3	4.1	2.3	3.2
94. seul	61.8	97.2	3.4	1.0	2.2
95. soie	40.9	56.0	4.1	3.6	3.8
96. sortir	53.9	85.3	3.7	2.3	3.0
97. sourire	10.4	50.5	5.7	3.8	4.7
98. suffire	9.8	10.2	5.7	5.7	5.7
99. surtout	19.7	69.7	5.1	3.0	4.0
100. tard	83.4	94.5	2.4	1.4	1.9
101. tête	75.4	98.2	2.8	0.7	1.7
102. toujours	81.6	94.5	2.5	1.4	1.9
103. traduire	30.7	68.8	4.5	3.1	3.8
104. troisième	58.2	81.7	3.5	2.5	3.0
105. un	98.0	96.3	0.8	1.2	1.0
106. vendre	75.2	86.2	2.8	2.2	2.5
107. vers	41.1	79.8	4.1	2.6	3.3
108. ville	87.0	90.8	2.1	1.8	1.9
109. visite	79.1	92.7	2.6	1.6	2.1
110. voilà	69.5	76.1	3.0	2.8	2.9
111. votre	89.1	94.5	2.0	1.4	1.7
112. vrai	70.9	93.5	3.0	1.6	2.3

TABLE II—CONTINUED

List 4	Percents Correct		Year Scale Values		General Scale Values
	I	II	I	II	
1. afin que	22.0	23.2	4.9	4.9	4.9
2. aller	98.8	100.0	0.4	0.0	0.2
3. an	52.9	70.4	3.7	3.0	3.3
4. apercevoir	74.7	92.0	2.8	1.7	2.2
5. apprendre	29.0	52.8	4.6	3.7	4.1
6. arrêter	64.9	93.6	3.2	1.5	2.3
7. aujourd'hui	92.9	98.4	1.6	0.6	1.1
8. autour	36.5	78.4	4.3	2.6	3.4
9. avec	98.4	100.0	0.6	0.0	0.3
10. besoin	51.0	78.4	3.8	2.6	3.2
11. blanc	96.7	100.0	1.1	0.0	0.5
12. bon	98.0	97.6	0.8	0.9	0.8
13. ça	41.7	76.8	4.1	2.7	3.4
14. ce	90.9	96.8	1.8	1.1	1.4
15. cent	81.5	97.6	2.5	0.9	1.7
16. chaise	72.4	92.0	2.9	1.7	2.3
17. chaque	53.7	90.4	3.7	1.9	2.8
18. chercher	76.1	95.2	2.8	1.3	2.0
19. choisir	84.4	97.6	2.3	0.9	1.6
20. clair	71.8	96.8	2.9	1.1	2.0
21. comme	74.7	95.2	2.8	1.3	2.0
22. conduire	46.3	83.6	3.9	2.4	3.1
23. corps	72.4	93.6	2.9	1.5	2.2
24. couper	45.7	79.2	4.0	2.6	3.3
25. couvrir	48.9	88.8	3.8	2.0	2.9
26. dame	92.9	95.2	1.6	1.3	1.4
27. défendre	92.1	97.6	1.7	0.9	1.3
28. demander	98.0	99.2	0.8	0.2	0.5
29. derrière	58.3	76.0	3.5	2.8	3.1
30. deuxième	68.3	84.4	3.1	2.3	2.7
31. difficile	94.6	97.6	1.4	0.9	1.1
32. doux	36.1	72.8	4.3	2.9	3.6
33. écrire	80.5	99.2	2.5	0.2	1.3
34. en (prep.)	56.6	85.6	3.6	2.2	2.9
35. entendre	59.5	84.8	3.4	2.2	2.8
36. espérer	56.8	92.0	3.5	1.7	2.6
37. faim	50.2	62.0	3.8	3.3	3.5
38. femme	92.1	97.6	1.7	0.9	1.3

TABLE II—CONTINUED

List 4—Continued	Percents Correct		Year Scale Values		General Scale Values
	I	II	I	II	
39. fille	99.0	100.0	0.3	0.0	0.1
40. fleur	88.4	95.2	2.0	1.3	1.6
41. frais	37.1	72.0	4.3	2.9	3.6
42. froid	87.8	100.0	2.1	0.0	1.0
43. gauche	59.5	82.0	3.4	2.4	2.9
44. gros	88.0	96.0	2.1	1.2	1.6
45. heure	90.3	98.4	1.9	0.6	1.2
46. histoire	98.6	98.4	0.5	0.6	0.5
47. il	97.8	98.4	0.8	0.6	0.7
48. je	98.8	98.4	0.4	0.6	0.5
49. jour	96.3	98.4	1.2	0.6	0.9
50. là	90.0	92.2	1.9	1.7	1.8
51. le (pron.)	73.4	76.0	2.9	2.8	2.8
52. lettre	96.3	100.0	1.2	0.0	0.6
53. ligne	42.7	87.2	4.1	2.1	3.1
54. loi	12.7	38.4	5.5	4.2	4.8
55. lui	90.0	96.0	1.9	1.2	1.5
56. maintenant	95.9	99.2	1.2	0.2	0.7
57. mal (adv.)	64.3	86.4	3.3	2.2	2.8
58. mauvais	79.4	79.2	2.6	2.6	2.6
59. mener	21.4	46.4	5.0	3.9	4.4
60. midi	60.8	86.4	3.4	2.2	2.8
61. moi	88.8	92.0	2.0	1.7	1.8
62. mon	88.0	89.4	2.1	1.9	2.0
63. montre	46.3	59.2	3.9	3.5	3.7
64. mot	64.9	98.4	3.2	0.6	1.9
65. neuf (no.)	87.6	95.2	2.1	1.3	1.7
66. non	92.9	97.6	1.6	0.9	1.2
67. nouveau	82.8	96.8	2.4	1.1	1.7
68. œil	61.6	88.8	3.4	2.0	2.7
69. or	46.3	73.6	3.9	2.9	3.4
70. où	85.9	88.0	2.2	2.1	2.1
71. ouvert	58.5	88.0	3.3	2.1	2.7
72. par	63.9	91.4	3.3	1.8	2.5
73. parmi	24.9	62.0	4.8	3.3	2.0
74. pas (n)	16.6	59.2	5.2	3.5	4.3
75. pays	67.4	96.8	3.1	1.1	2.1

TABLE II—CONTINUED

List 4—Continued	Percents Correct		Year Scale Values		General Scale Values
	I	II	I	II	
76. perdre	55.6	84.4	3.6	2.3	2.9
77. petit (adj.)	99.6	99.2	0.0	0.2	0.1
78. plaire	35.3	80.0	4.4	2.6	3.5
79. plus	90.7	96.0	1.8	1.2	1.5
80. porte	75.1	85.6	2.8	2.2	2.5
81. pourvu que	7.7	20.0	5.9	5.0	5.4
82. près de	66.4	87.2	3.2	2.1	2.6
83. professeur	96.6	98.4	1.1	0.6	0.8
84. quart	52.1	68.0	3.7	3.1	3.4
85. question	98.6	98.4	0.5	0.6	0.5
86. quoi	57.7	83.6	3.5	2.4	2.9
87. règle	36.5	66.0	4.3	3.2	3.7
88. remplir	37.6	67.2	4.3	3.1	3.7
89. réussir	21.8	39.2	5.0	4.2	4.6
90. rire	47.7	88.0	3.9	2.1	3.0
91. salle	84.6	92.2	2.3	1.7	2.0
92. se	45.9	55.2	4.0	3.6	3.8
93. sentir	14.1	64.0	5.4	3.3	4.3
94. seulement	54.1	79.2	3.6	2.6	3.1
95. soif	19.5	34.4	5.1	4.4	4.7
96. son (poss. adj.)	91.3	96.8	1.8	1.1	1.4
97. sous	53.3	83.2	3.7	2.4	3.0
98. suite	23.0	32.0	4.9	4.5	4.7
99. table	98.8	98.4	0.4	0.6	0.5
100. te	56.0	73.6	3.6	2.9	3.2
101. toi	51.6	66.0	3.7	3.2	3.4
102. tort	15.8	50.4	5.3	3.8	4.5
103. travail	67.6	91.2	3.1	1.8	2.4
104. trop	59.1	76.0	3.5	2.8	3.1
105. utile	57.0	72.0	3.5	2.9	3.2
106. venir	73.0	96.8	2.9	1.1	2.0
107. vert	71.0	93.6	3.0	1.5	2.2
108. vin	64.9	95.2	3.2	1.3	2.2
109. vite	51.5	83.2	3.7	2.4	3.0
110. voir	83.6	99.2	2.4	0.2	1.3
111. vouloir	71.6	93.6	3.0	1.5	2.2
112. y	77.6	86.4	2.7	2.2	2.4

From the total list of 448 words a selection was made by eliminating all those words which were never missed by either first- or second-year pupils, those whose meanings are readily inferred from the English correspondents, those which showed inversions, being translated correctly more often by first-year pupils than by second-year pupils, and those words which experience had shown were troublesome for one reason or another to score. This reduced the list to 240 words which were grouped into four sixty-word standard vocabulary tests of equal difficulty. These tests with the weights to be given each word and the sum of the scale values are as follows:

FRENCH VOCABULARY TESTS[2]

A		B		C		D	
aujourd'hui	1.1	acheter	1.2	soleil	1.6	autre	1.4
chose	1.7	chien	1.7	alors	1.7	choisir	1.6
nouveau	1.7	rue	1.7	tête	1.7	vingt	1.7
porter	1.9	tard	1.9	crier	1.9	sans	1.8
hiver	1.9	toujours	1.9	parmi	2.0	chercher	2.0
onze	2.0	salle	2.0	frapper	2.0	comprendre	2.0
venir	2.0	même	2.0	pays	2.1	fort	2.1
vieux	2.1	vivre	2.1	apercevoir	2.2	guerre	2.1
mille	2.1	où	2.1	partir	2.2	morceau	2.2
vert	2.2	permettre	2.2	neige	2.2	mort	2.2
vouloir	2.2	bois	2.2	arrêter	2.3	vin	2.2
année	2.3	hier	2.3	obéir	2.3	lever	2.2
haut	2.3	soir	2.3	demain	2.3	jouer	2.3
semaine	2.3	connaître	2.3	comment	2.4	combien	2.3
meilleur	2.4	ouvrir	2.4	ni	2.4	promener	2.4
apporter	2.4	travail	2.4	savoir	2.4	quinze	2.4
dormir	2.4	cela	2.5	champ	2.4	remercier	2.4
vie	2.4	noir	2.5	ciel	2.5	ne	2.4
lire	2.4	sucre	2.5	rien	2.5	après	2.5
devant	2.5	vendre	2.5	viande	2.5	pied	2.5
lait	2.5	coucher	2.6	billet	2.5	par	2.5
vache	2.5	droit	2.6	mauvais	2.6	porte	2.5
déjà	2.6	près de	2.6	pouvoir	2.6	avant	2.6
quitter	2.6	chemin	2.7	deuxième	2.7	espérer	2.6
bras	2.7	œil	2.7	mourir	2.7	feu	2.7
on	2.7	voix	2.7	courir	2.7	chaque	2.8
ouvert	2.7	entendre	2.8	côté	2.8	chaud	2.8
jeter	2.8	mer	2.8	midi	2.8	moins	2.8

[2] These tests can be procured from the writer, University of Wisconsin, Madison, Wisconsin.

FRENCH VOCABULARY TESTS—CONTINUED

A		B		C		D	
mois	2.9	dernier	2.8	voici	2.8	couvrir	2.9
oublier	2.9	perdre	2.9	quoi	2.9	tante	2.9
gauche	2.9	attendre	2.9	vent	2.9	voilà	2.9
mieux	2.9	laisser	2.9	assez	3.0	ceci	3.0
rester	3.0	printemps	3.0	car	3.0	fin	3.0
rire	3.0	sortir	3.0	gens	3.0	souvent	3.0
chacun	3.1	vite	3.0	sous	3.0	troisième	3.0
seulement	3.1	bruit	3.1	tenir	3.0	devoir	3.0
conduire	3.1	pendant	3.1	enfin	3.1	cinquante	3.1
quelque	3.2	trop	3.1	maître	3.1	ligne	3.1
rendre	3.2	mettre	3.2	derrière	3.1	besoin	3.2
sembler	3.2	ton	3.2	lit	3.2	lequel	3.2
gagner	3.3	entre	3.3	utile	3.2	voiture	3.2
ou	3.3	partie	3.3	couper	3.3	an	3.3
cependant	3.4	ça	3.4	plusieurs	3.3	vers	3.3
plein	3.4	or	3.4	boire	3.4	autour	3.4
ordonner	3.5	faim	3.5	puis	3.4	suivre	3.4
plaire	3.5	presque	3.5	envoyer	3.5	depuis	3.6
ainsi	3.6	peur	3.6	croire	3.6	doux	3.6
rentrer	3.6	fer	3.6	frais	3.6	montre	3.7
oreille	3.7	pleurer	3.7	bientôt	3.7	règle	3.7
chez	3.8	verre	3.7	prêt	3.7	remplir	3.7
dont	3.8	fou, fol	3.8	sauver	3.7	traduire	3.8
lorsque	4.0	journée	4.0	soie	3.8	bout	3.9
soulier	4.0	surtout	4.0	tant	3.9	apprendre	4.1
paraître	4.2	loin	4.2	craindre	4.2	couteau	4.1
déjeuner	4.2	falloir	4.3	sentir	4.3	coup	4.1
aussitôt	4.5	court	4.4	mener	4.4	devenir	4.4
moindre	4.5	tort	4.5	réussir	4.6	nul	4.6
autant	4.8	sourire	4.7	soif	4.7	minuit	4.7
lien	5.2	quoique	5.0	afin que	4.9	loi	4.8
valoir	5.4	appartenir	5.3	pourvu que	5.3	suffire	5.7
	179.6		179.6		179.6		179.4

Tables III and IV give the results with these tests for sixteen schools in terms of the number right and the sum of the scale values.

TABLE III. NUMBER CORRECT

Test	Years of Instruction in French				
	0.5	1.0	1.5	2.0	3.0
A	22.1	32.4	45.4	47.6	53.3
B	25.7	32.6	44.9	47.8	53.5
C	22.4	31.6	42.2	46.6	52.3
D	29.3	32.7	43.7	47.0	52.6
Average	24.9	32.3	44.1	47.2	52.9

TABLE IV. SUM OF SCALE VALUES

Test	Years of Instruction in French				
	0.5	1.0	1.5	2.0	3.0
A	59.1	87.1	124.6	132.8	154.7
B	69.3	87.2	124.5	134.8	156.1
C	60.5	86.4	114.2	131.1	150.7
D	77.0	87.7	121.3	132.3	151.1
Average	66.5	87.1	121.1	132.7	153.1

Tentative standard scores for the four vocabulary tests are, therefore, as given in Table V.

TABLE V. AVERAGE OF SCORES FOR THE FOUR TESTS

	Years of Instruction in French				
	0.5	1.0	1.5	2.0	3.0
Number correct	25	32	44	47	53
Sum of scale values	66	87	121	133	153

When the tests are given the score may simply be represented by the number of words correct. It is more accurate, however,

to give the weight or scale value for each correct word and take the sum of the scale values as the score.

Sentence Tests

Seventy-five sentences, containing no words not in the list of 448 common to all the first year, were selected or constructed and arranged in five sentence tests of as nearly equal difficulty as could be estimated. A large number of sentences were selected from exercise and composition books used with first- and second-year pupils. The sentences were selected in this way in order that tests which might be developed would be based upon the sort of material the pupils are usually called upon to work with, and, moreover, such material would lend itself readily to the development of construction or grammar tests.

Table VI gives the results obtained from these seventy-five sentences in eighteen schools distributed widely over the country. As before the average and median scores are given and the number of pupils. While the tests were given to half-year, year-and-a-half, and three-year pupils, the results are not used in the calculation of scale values.

TABLE VI. RESULTS FOR THE SEVENTY-FIVE SENTENCES ARE ARRANGED IN THE ORIGINAL FIVE SENTENCE TESTS. ALL SCHOOLS COMBINED

Tests	First-Year Pupils			Second-Year Pupils		
	No. of Pupils	Averages	Medians	No. of Pupils	Averages	Medians
I	488	7.8	8.3	88	10.7	11.8
II	474	7.0	7.4	103	10.4	11.5
III	414	7.0	7.5	116	10.5	10.9
IV	507	7.6	7.9	132	10.0	10.8
V	408	7.8	8.9	125	12.3	12.5

Table VII gives (a) the percent of times each of the seventy-five sentences is correctly translated by first- and second-year pupils; (b) the year scale values (derived by the same method as was used with vocabularies described above); and (c) the general scale value or weight to be assigned to each sentence.

TABLE VII. THE SEVENTY-FIVE SENTENCES GIVEN TO FIRST- AND SECOND-YEAR PUPILS

	Percents Correct		Year Scale Values		General Scale Values
	I	II	I	II	
1. J'ai un frère	96.3	96.6	1.2	1.1	1.1
2. Ce livre est petit	97.5	100.0	0.9	0.0	0.4
3. Il n'a pas d'argent	90.4	89.8	1.9	1.9	1.9
4. La jeune fille a les yeux bleus	95.1	95.5	1.4	1.3	1.3
5. J'y suis allé ce matin	43.4	65.8	4.0	3.2	3.6
6. Il vient de vendre ce cheval quatre cents francs	11.1	37.5	5.6	4.3	4.9
7. On trouve que la vie est chére dans les grandes villes	73.9	89.8	2.9	1.9	2.4
8. Il mangea très peu quoiqu'il eût faim	25.8	51.1	4.8	3.8	4.3
9. Reste avec nous où il fait chaud	53.5	81.8	3.7	2.5	3.1
10. Il l'a fait sans qu'on le lui ait dit	19.7	44.3	5.1	4.0	4.5
11. J'ai dû partir sans vous attendre	17.0	48.8	5.2	3.8	4.5
12. Si vous connaissez ce monsieur qui vient d'arriver, dites-moi son nom	25.4	56.8	4.8	3.5	4.1
13. Elles se parlent l'une à l'autre	48.7	67.0	3.8	3.1	3.4
14. Je l'ai vu il y a quelques heures	34.2	68.2	4.4	3.1	3.7
15. Aussitôt qu'il a eu fini son travail, ses amis sont arrivés	52.9	78.4	3.7	2.6	3.1
16. C'est vrai	85.4	94.2	2.2	1.5	1.8
17. Le chien est l'ami de l'homme	95.3	97.1	1.3	1.0	1.1
18. Il est difficile de faire cela	86.0	90.3	2.2	1.9	2.0
19. Quel est le livre que vous avez à la main?	93.2	98.1	1.6	0.7	1.1
20. Je ne le dirai pas à vous	48.6	82.5	3.9	2.4	3.1
21. C'est pourquoi nous ne sommes pas allés sur l'eau	50.7	79.6	3.8	2.6	3.2
22. J'ai déjà été le voir	61.3	85.4	3.4	2.2	2.8
23. Il devait arriver hier	29.6	63.1	4.6	3.3	3.9
24. Qu'est-ce que ça veut dire?	10.0	31.1	5.7	4.5	5.1
25. Je ne crains pas qu'il ne soit pas venu	15.2	39.8	5.3	4.2	4.7
26. Il y a un mois que je le connais	10.4	32.0	5.7	4.5	5.1
27. Il a de son côté quelque chose à dire, il me semble	23.9	68.0	4.9	3.1	4.0
28. S'il a le temps, dites-lui d'y aller	54.1	83.5	3.6	2.4	3.0
29. Qui pourra le sauver si ce n'est ses amis	24.5	54.4	4.8	3.6	4.2
30. Et quand cela aurait été vrai, je ne l'aurais pas dit	15.9	40.8	5.3	4.1	4.7

TABLE VII—CONTINUED

	Percents Correct		Year Scale Values		General Scale Values
	I	II	I	II	
31. Il est heureux	91.6	93.9	1.8	1.5	1.6
32. Le vent et la neige sont arrivés	86.9	92.2	2.1	1.7	1.9
33. Oui, maintenant il y est	75.4	87.1	2.8	2.1	2.4
34. Je suis ici depuis trois jours	34.8	67.2	4.4	3.1	3.7
35. Je n'ai parlé à personne	83.6	92.2	2.4	1.7	2.0
36. Vous écrirez ce qu'il vous plaira	21.9	64.7	5.0	3.2	4.1
37. Il me l'a dit lui-même	50.2	82.7	3.9	2.4	3.1
38. J'ai peur qu'il ne vienne	6.3	12.9	6.1	5.5	5.8
39. Je vous donnerai ce dont vous avez besoin	60.1	83.6	3.4	2.4	2.9
40. Je ne sais pas pourquoi nos amis ne sont pas encore revenus	61.1	87.1	3.4	2.1	2.7
41. Je resterai ici pour qu'ils puissent me trouver	50.2	80.2	3.8	2.5	3.1
42. Si je m'en vais, ce n'est pas que je veuille vous quitter	21.5	58.6	5.0	3.5	4.2
43. Ils ont dit qu'ils me feraient savoir quand ils arriveraient	30.0	73.3	4.6	2.9	3.7
44. Le soleil se lève de très bon matin et couche si tard que, à huit heures, il fait encore jour	20.5	71.5	5.0	3.0	4.0
45. Peu s'en faut qu'il ne le fasse	9.4	6.0	5.8	6.1	5.9
46. Qui est là?	92.3	95.4	1.7	1.3	1.5
47. La nuit tombe	83.0	89.4	2.4	1.9	2.1
48. Je suis votre ami, n'est-ce pas?	97.4	99.2	0.9	0.2	0.5
49. Il est difficile de lui plaire	61.3	81.1	3.4	2.5	2.9
50. Est-ce qu'il n'est pas ici?	77.3	84.1	2.7	2.3	2.5
51. Cette montre vaut vingt francs	44.0	72.7	4.0	2.9	3.4
52. Il a perdu son chemin	51.9	84.9	3.8	2.3	3.0
53. Combien y a-t-il d'ici à votre maison?	29.2	33.3	4.6	4.4	4.5
54. Avez-vous jamais vu quelque chose d'aussi beau?	54.8	82.6	3.6	2.4	3.0
55. Je comprends maintenant que c'était mon devoir	20.9	59.1	5.0	3.5	4.2
56. Il s'y trouve de belles choses	9.1	3.8	5.8	6.4	6.1
57. On m'a dit que ce n'est pas vrai	73.8	90.9	2.9	1.8	2.3
58. Il viendra pourvu qu'il ne fasse pas froid	10.8	22.0	5.6	4.9	5.2
59. J'irai le trouver si vous n'y allez vous-même.	41.8	65.1	4.1	3.2	3.6
60. Faites semblant d'être content afin que les autres n'aient pas peur	9.7	40.9	5.7	4.1	4.9

TABLE VII—CONTINUED

	Percents Correct		Year Scale Values		General Scale Values
	I	II	I	II	
61. La mère aime l'enfant	84.6	96.0	2.3	1.2	1.7
62. Mon travail est fini	83.1	94.4	2.4	1.4	1.9
63. Ils sont à l'école	88.7	97.6	2.0	0.9	1.4
64. Tout le monde est ici aujourd'hui	88.5	96.8	2.0	1.1	1.5
65. Il doit arriver demain	33.3	73.6	4.4	2.9	3.6
66. Cet arbre est plus haut que le nôtre	67.2	91.2	3.1	1.8	2.4
67. Il paraît qu'il est mort	12.8	70.4	5.5	3.0	4.2
68. Voilà ce qui me plaît	62.2	88.0	3.3	2.1	2.7
69. Je cherche un pays où je puisse être libre	47.8	92.8	3.9	1.6	2.7
70. J'ai peu dormi cette nuit	41.7	66.4	4.1	3.2	3.6
71. Si vous étiez resté ici, j'aurais été content	38.2	73.6	4.2	2.9	3.5
72. Que ne lui avez-vous dit tout ce qui s'est passé?	15.7	46.4	5.3	3.9	4.6
73. Ne vous semble-t-il pas qu'il est fou?	23.5	73.6	4.9	2.9	3.9
74. Je ne veux pas qu'il y aille	38.1	80.8	4.2	2.5	3.4
75. Il leur a demandé de l'argent, mais ils ne lui ont rien donné	62.0	90.4	3.3	1.9	2.6

The sentences were then rearranged into six standard sentence tests of twelve sentences each of approximately equal difficulty. The tests are as follows:

FRENCH SENTENCE TEST I

1.	Je suis votre ami, n'est-ce pas?	0.5
2.	Quel est le livre que vous avez à la main?	1.1
3.	Tout le monde est ici aujourd'hui	1.5
4.	Il est difficile de faire cela	2.0
5.	Il leur a demandé de l'argent, mais ils ne lui ont rien donné	2.6
6.	Reste avec nous où il fait chaud	3.1
7.	J'y suis allé ce matin	3.6
8.	Vous écrirez ce qu'il vous plaira	4.1
9.	Que ne lui avez-vous dit tout ce qui s'est passé?	4.6
10.	Qu'est-ce que ça veut dire?	5.1
11.	J'ai peur qu'il ne vienne	5.8
12.	Il s'y trouve de belles choses	6.1
		40.1

FRENCH SENTENCE TEST II

1.	Le chien est l'ami de l'homme	1.1
2.	Qui est là?	1.5
3.	Le vent et la neige sont arrivés	1.9
4.	On m'a dit que ce n'est pas vrai	2.3
5.	Je cherche un pays où je puisse être libre	2.7
6.	Aussitôt qu'il a eu fini son travail, ses amis sont arrivés	3.1
7.	Si vous étiez resté ici, j'aurais été content	3.5
8.	Ne vous semble-t-il pas qu'il est fou?	3.9
9.	Il mangea très peu quoiqu'il eût faim	4.3
10.	Je ne crains pas qu'il ne soit pas venu	4.7
11.	Il y a un mois que je le connais	5.1
12.	Peu s'en faut qu'il ne le fasse	5.9
		40.0

FRENCH SENTENCE TEST III

1.	La jeune fille a les yeux bleus	1.3
2.	La mère aime l'enfant	1.7
3.	Je n'ai parlé à personne	2.0
4.	On trouve que la vie est chère dans les grandes villes	2.4
5.	J'ai déjà été le voir	2.8
6.	C'est pourquoi nous ne sommes pas allés sur l'eau	3.2
7.	J'irai le trouver si vous n'y allez vous-même	3.6
8.	Il a de son côté quelque chose à dire, il me semble	4.0
9.	Qui pourra le sauver si ce n'est ses amis	4.2
10.	Il l'a fait sans qu'on le lui ait dit	4.5
11.	Faites semblant d'être content afin que les autres n'aient pas peur	4.9
12.	Il viendra pourvu qu'il ne fasse pas froid	5.2
		39.8

FRENCH SENTENCE TEST IV

1.	Ils sont a l'école	1.4
2.	Mon travail est fini	1.9
3.	Cet arbre est plus haut que le nôtre	2.4
4.	Je ne sais pas pourquoi nos amis ne sont pas encore revenus	2.7
5.	S'il a le temps, dites-lui d'y aller	3.0
6.	Cette montre vaut vingt francs	3.4
7.	J'ai peu dormi cette nuit	3.6
8.	Il devait arriver hier	3.9
9.	Il paraît qu'il est mort	4.2
10.	J'ai dû partir sans vous attendre	4.5
11.	Et quand cela aurait été vrai, je ne l'aurais pas dit	4.7
12.	Il vient de vendre ce cheval quatre cents francs	4.9
		40.6

FRENCH SENTENCE TEST V

1.	C'est vrai	1.8
2.	La nuit tombe	2.1
3.	Oui, maintenant il y est	2.4
4.	Voilà ce qui me plaît	2.7
5.	Je vous donnerai ce dont vous avez besoin	2.9
6.	Je resterai ici pour qu'ils puissent me trouver	3.1
7.	Je ne veux pas qu'il y aille	3.4
8.	Ils ont dit qu'ils me feraient savoir quand ils arriveraient	3.7
9.	Le soleil se lève de très bon matin et se couche si tard que, à huit heures, il fait encore jour	4.0
10.	Si vous connaissez ce monsieur qui vient d'arriver, dites-moi son nom	4.1
11.	Si je m'en vais, ce n'est pas que je veuille vous quitter	4.2
12.	Combien y a-t-il d'ici à votre maison?	4.5
		38.9

FRENCH SENTENCE TEST VI

1.	Il n'a pas d'argent	1.9
2.	Est-ce qu'il n'est pas ici?	2.5
3.	Il est difficile de lui plaire	2.9
4.	Avez-vous jamais vu quelque chose d'aussi beau?	3.0
5.	Il a perdu son chemin	3.0
6.	Je ne le dirai pas à vous	3.1
7.	Il me l'a dit lui-même	3.1
8.	Elles se parlent l'une à l'autre	3.4
9.	Il doit arriver demain	3.6
10.	Je l'ai vu il y a quelques heures	3.7
11.	Je suis ici depuis trois jours	3.7
12.	Je comprends maintenant que c'était mon devoir	4.2
		38.1

Table VIII gives the number correct with each of these tests and the sum of the scale values for one year, a year and one-half, two year, and three year pupils.

TABLE VIII. NUMBER CORRECT

Test	Years of Instruction in French			
	1.0	1.5	2.0	3.0
I	5.9	7.4	7.7	9.7
II	5.8	7.4	8.5	10.0
III	5.4	7.2	8.2	8.6
IV	5.2	7.3	8.7	10.4
V	5.7	7.5	9.0	10.5
VI	6.1	8.3	9.2	11.0
Average	5.7	7.5	8.5	10.0

TABLE IX. SUM OF SCALE VALUES

Test	Years of Instruction in French			
	1.0	1.5	2.0	3.0
I	12.7	17.3	18.8	28.1
II	14.1	19.6	23.2	30.5
III	14.0	20.0	23.6	29.0
IV	14.2	21.3	26.8	34.1
V	16.2	22.5	27.5	32.6
VI	18.1	26.0	30.0	33.7
Average	14.9	21.1	25.0	31.3

Tentative standard scores, disregarding the minor differences between the tests, would, therefore, be as follows:

TABLE X. AVERAGE OF SCORES FOR THE SIX TESTS

	Years of Instruction in French			
	1.0	1.5	2.0	3.0
Number correct	5.5	7.5	8.5	10.0
Sum of scale values	14.0	21.0	24.0	31.0

These scores are very reliable for the one year and two year pupils but less so for the year and a half and three year pupils.

THE OTIS GROUP INTELLIGENCE SCALE APPLIED TO THE ELEMENTARY-SCHOOL GRADUATING CLASSES OF OAKLAND, CALIFORNIA

VIRGIL E. DICKSON AND JOHN K. NORTON
Bureau of Research and Guidance, Public Schools, Oakland, California

Both individual and group intelligence tests have been extensively used in the public schools of Oakland, California for a number of years. Their value as administrative tools is well established. They have contributed to a better understanding of many school problems such as retardation, elimination, and proper curriculum modification; and they have resulted in administrative adjustments leading to greater efficiency.

This article presents certain data from tests of 1,043 day elementary-school pupils who completed the eighth grade in Oakland, January, 1920. This includes all those graduated at that time except a small number who were absent from school the day the tests were given. The Otis Group Intelligence Scale, as abbreviated for use in Oakland, was employed in making the tests. The abbreviation omits tests 1, 6, 9, and 10. It includes tests 2, 3, 4, 5, 7, and 8 without modification as to directions for giving, timing, or scoring, except that tests 4 and 5 are given double value in the final score in order that they may contribute a greater amount to the final test result. The highest possible score with the abbreviated test is 175.

In order that the six-test scores obtained at Oakland may be compared with the scores arising from the use of the complete test, Dr. Arthur S. Otis, the deviser of the test, has furnished Table I. By its use a ten-test score may be transmuted into a six-test score.

The tens and hundreds digits of the six-test score are printed in italics in the line across the top; the units digits of the six-test score are printed in italics in the column at the left. To find the ten-test score corresponding to a known six-test score find in the body of the table the number under the tens digit (or tens and hundreds digits combined) and opposite the units digit. For ex-

TABLE I. A TABLE FOR TRANSMUTING TEN-TEST SCORES INTO SIX-TEST SCORES, OMITTING TESTS 1, 6, 9, AND 10. (OTIS GROUP INTELLIGENCE SCALE)

Units Digits of Six-test Score	Tens and Hundreds Digits of Six-test Score												
	0	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*
1	10	26	42	58	74	91	107	123	140	154	167	181	195
2	11	27	43	60	76	92	109	125	141	155	168	182	196
3	13	29	45	61	78	94	110	127	143	157	170	184	198
4	14	30	46	63	79	96	112	128	144	158	171	185	199
5	16	32	48	64	81	97	114	130	146	160	173	187	201
6	18	34	50	66	82	99	115	132	147	161	174	188	202
7	19	35	51	68	84	100	117	133	148	162	175	189	203
8	21	37	53	69	86	102	118	135	150	163	177	191	205
9	22	38	55	71	87	104	120	136	151	164	178	192	206
0	24	40	56	73	89	105	122	138	153	166	180	194	208

ample the following six-test scores correspond to the indicated ten-test scores:

Six-test Score	Ten-test Score
45	81
125	201
22	43
10	24 (40)
4	14

The tests were made in the school month just preceding the date of graduation. At the time of making the test, the following supplementary data were obtained for each child: name, age in years and months, date and hour of the test, and teachers' ratings in school work. Teachers were asked to rate pupils for general quality of school work on a scale of seven points as follows: (1) very superior; (2) superior; (3) slightly above average; (4) average; (5) slightly below average; (6) inferior; (7) very inferior.

When several teachers rated a child in school work, the average of their ratings was taken. If this resulted in a fraction, the child was assigned the rating of the nearest whole number. When the fraction was an even half, the rating was given the value of the

TABLE II. DISTRIBUTIONS OF SCORES MADE BY ELEMENTARY-SCHOOL GRADUATING CLASSES IN 29 SCHOOLS
OTIS GROUP INTELLIGENCE TEST—OAKLAND ABBREVIATION[a]

Score	School																												Total
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	
150	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	1	....	....	....	1	1	2
145	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	1	....	....	....	....	....	....	2	....	3	6
140	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	1	....	....	....	....	3	4	8
135	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	2	....	1	1	1	....	1	....	3	2	11
130	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	1	....	....	3	....	....	1	6	6	21
125	....	....	....	....	....	....	....	....	....	....	....	....	1	1	....	....	1	....	....	....	1	....	1	4	1	2	1	6	2	27
120	....	....	....	....	....	1	....	....	1	....	....	....	2	....	....	1	1	2	....	1	2	....	....	3	2	3	3	11	7	41
115	....	....	....	1	....	1	....	....	1	....	1	....	....	1	1	....	....	....	1	....	....	3	1	3	2	....	1	11	9	44
110	....	1	....	1	....	....	....	....	....	2	....	....	2	3	2	....	1	1	....	1	2	1	1	4	2	....	1	11	7	43
105	....	....	....	....	....	....	1	....	1	....	1	1	1	1	1	....	1	1	1	1	2	3	....	4	3	....	2	11	7	63
100	....	....	....	....	2	2	2	....	2	1	2	2	....	1	2	1	3	2	1	2	2	5	4	3	8	2	....	9	5	57
95	....	....	1	2	....	3	3	....	1	3	....	1	....	3	2	....	2	2	2	2	1	....	4	3	2	3	3	11	3	97
90	1	....	2	....	3	4	3	1	2	1	2	5	4	7	....	1	5	4	3	2	1	3	3	6	8	2	2	18	4	76
85	....	1	....	....	3	6	1	....	3	....	3	2	2	....	1	3	5	2	3	....	2	1	6	6	6	2	....	12	3	76
80	1	1	1	2	2	1	6	1	3	....	1	4	3	1	5	3	4	5	4	2	3	1	....	4	2	2	2	10	2	74
75	1	1	....	1	....	3	3	2	....	1	4	4	6	6	2	....	5	4	2	2	3	3	3	2	4	2	3	6	1	85
70	....	2	2	....	6	2	6	1	2	1	3	6	7	4	1	3	3	3	2	2	3	3	3	3	5	2	2	7	4	70
65	2	2	2	....	6	7	3	1	1	....	4	2	4	4	2	2	1	3	3	3	1	1	3	6	4	....	1	3	1	61
60	1	2	....	2	2	2	2	1	3	1	1	5	3	1	4	1	5	1	3	1	2	3	....	7	2	....	2	5	2	61
55	1	5	....	2	3	5	3	....	4	....	1	1	2	4	4	1	1	2	1	....	1	3	1	2	1	1	....	3	1	53
50	....	2	4	....	4	4	1	1	2	2	2	1	....	....	1	....	2	1	....	2	1	4	....	1	1	....	....	1	1	39
45	2	4	3	2	1	3	0	1	3	2	2	1	1	1	....	1	1	....	....	....	....	....	....	....	....	....	1	....	1	36
40	4	1	....	2	2	1	3	....	1	1	....	....	....	....	....	....	....	....	....	....	....	....	....	1	....	....	....	....	....	17
35	1	....	1	....	1	1	1	....	....	1	....	....	....	....	....	....	....	....	....	1	....	....	1	1	3	....	....	1	....	10
30	....	1	1	....	....	2	2	....	....	....	....	....	....	....	....	....	1	....	....	....	....	....	1	....	....	....	....	....	....	11
25	....	1	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	2	1	....	....	....	....	....	....	....	....	....	....	4
20	1	....	....	....	2	1	2	....	....	....	....	1	....	....	....	....	....	1	1	....	....	....	....	....	....	....	....	....	....	8
15	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	1	1
10	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	1
Totals.....	15	24	17	15	37	49	48	9	30	16	27	36	38	38	28	17	42	34	24	25	26	41	33	67	51	22	26	138	70	1,043
Medians....	48	58	54	64	68	69	71	72	72	75	76	76	77	79	80	81	82	82	82	84	85	87	88	88	90	95	95	96	109	83

[a] Read table as follows: In the eighth-grade class in school 1, one pupil made a score of 90-94, one scored 80-84, etc. There was a total of 15 graduating pupils in school 1, and they made a median score of 48. Similar data are given for the 28 other schools graduating eighth-grade pupils.

whole number farthest removed from the rating of 4. That is 3.5 was tabulated as 3, 4.5 as 5, etc. Succeeding paragraphs present and briefly point out the significance of the data resulting from the test and the collection of the supplementary information.

Table II gives distributions of scores made by graduating classes of the 29 schools of the city. A wide variation in capacity as measured by the test is shown both by different high eighth grades and by different individuals in the same class. Individual scores range from 14 to 152 points and class medians from 48 to 109. The significance of these differences can be better understood by consulting Figure 1. This compares the ranges of the middle 50 percents for four different groups: (1) the elementary graduating classes of the two schools making the highest medians; (2) a high ninth-grade class of average ability tested at about the same time as the elementary graduating classes; (3) the elementary graduating classes in the three schools making the lowest median scores; (4) two high sixth-grade classes recently tested.

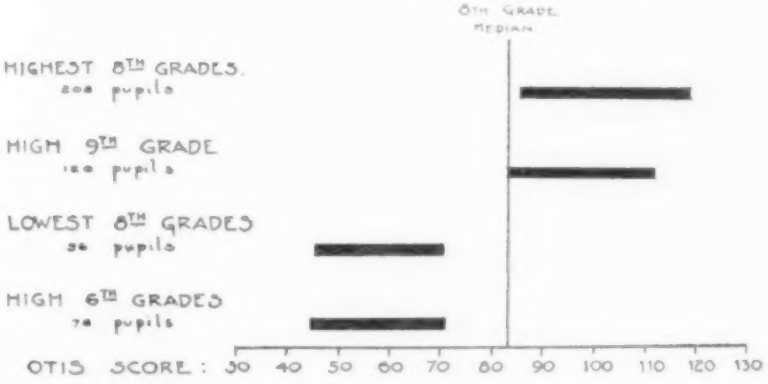

FIGURE 1. RANGE OF OTIS TEST SCORES
MIDDLE FIFTY PERCENT

It is evident that ability as measured by the test in the eighth-grade classes in some of the schools is equal to that in a freshman class which has survived a year in high school and is just ready to begin sophomore work; while the ability found in other eighth grades is no better than that of groups of children who have not yet completed the sixth grade. If the test means anything in

terms of ability to do school work, certainly the actual accomplishments of the different classes graduating in January, 1920 must have varied widely.

These data show how important it is to know the character of the human material with which a teacher is working, before a judgment is passed on the results obtained. A poor teacher working with an eighth-grade class with the capacity of high-school freshmen might very easily get better results than a good teacher working with an eighth-grade class with a capacity not above that of the ordinary sixth grade. Teachers are too often judged wholly on a basis of results, whereas the important factor, the learning capacity of children, is given but little thought.

The necessity of making modifications in the curriculum to meet the varying abilities of different classes of the same grade designation, and of different individuals in the same class, is another point emphasized by the data of Table II.

Modifications certainly must be made either in the character of the curriculum or in the time given by different classes of the same grade to any particular block of subject-matter if proper consideration is to be given to the wide differences in capacity which the test reveals in grades of the same designation.

Table III gives the data concerning the relationship between test scores and ratings in school work. There is a positive correlation of 0.538 (Pearson formula). Even were the test perfect in its measurement of mental capacity, we should not expect a complete correlation between test scores and school work. Such factors as the following all tend to destroy correlation:

1. All children do not put forth equal effort; some of good native ability may do poor work due to lack of application, and some of poor native ability may compensate by unusual effort.

2. Repetition of grades sometimes brings up the school work of those of poor native ability.

3. Desirable or undesirable personal characteristics may unduly influence a teacher in her estimate of a child's work.

4. Different classes of the same grade vary so widely in general ability that the same score may represent high or low ability, according to the class in which it occurs.

5. In a small number of cases the test score is probably unduly influenced by extreme nervousness, illness, or other interfering factors.

TABLE III. OTIS SCALE SCORE AND RATING IN SCHOOL WORK[a]

Score	School Work Rating							Totals
	7	6	5	4	3	2	1	
150	.....	.....	.....	.....	.....	1	.....	1
145	.....	.....	.....	.....	.....	1	1	2
140	.....	.....	.....	.....	2	.....	4	6
135	.....	.....	.....	.....	1	3	4	8
130	.....	.....	.....	4	1	3	2	10
125	.....	.....	1	4	9	6	1	21
120	.....	1	.....	5	10	10	1	27
115	.....	.....	3	7	21	8	2	41
110	.....	.....	.....	10	14	16	4	44
105	.....	1	3	11	21	7	.....	43
100	.....	1	12	25	17	8	.....	63
95	.....	1	6	29	12	7	2	57
90	.....	1	7	46	26	13	2	95
85	.....	1	9	34	23	9	.....	76
80	.....	1	13	32	15	12	.....	73
75	1	7	11	31	10	7	2	69
70	1	7	23	36	11	4	.....	82
65	4	7	23	20	11	3	2	70
60	2	8	12	28	11	.....	.....	61
55	1	8	21	11	6	1	.....	48
50	4	10	10	10	4	1	.....	39
45	2	9	6	5	9	.....	.....	31
40	1	4	6	3	.....	.....	.....	14
35	.....	6	1	2	.....	.....	.....	9
30	2	3	4	.....	1	.....	.....	10
25	.....	2	2	.....	.....	.....	.....	4
20	1	1	3	2	.....	.....	.....	7
15–19	.....	.....	1	.....	.....	.....	.....	1
10–14	.....	1	.....	.....	.....	.....	.....	1
Totals	19	80	177	355	235	120	27	1013[b]
Median Scores	54	57	70	85	93	102	119	84

r=O. 538.

[a] Read table as follows: Of the pupils rated 7 (very inferior) in school work one scored between 75 and 79; one scored between 70 and 74; etc. Of those rated 6 (inferior), one scored between 120 and 124; one between 105 and 109; etc.

[b] Thirty pupils included in Table II are omitted as they had not been in the Oakland schools a sufficient time to make it possible for teachers to rate them in school work.

6. The degree of homogeneity of the group tested is an important consideration in estimating the value of a correlation figure. The group dealt with here, high eight-grade children, is probably as homogeneous as is found in any grade.

But in spite of all these factors that tend to destroy correlation, the test score gives a good indication of the character of school work a child will do. Table IV, worked out on the percentage basis, is more conclusive on this point than the correlation figure. Of the fifty children composing the 5 percent making the highest test scores, none were rated 7 in school work, none were rated 6, 2 percent were rated 5 or below average, and the rest were rated average or above. School ratings for the other percentage groups are also given. These ratings show conclusively that the test score is a valuable indication of the type of work children do.

The data contained in Table V show the relation between chronological age and test scores for the 1,043 children. The younger children in the grade generally made the higher scores in the test. Those who were twelve years old made a median score of 110, and there was a marked drop in the median of each higher age group. In short, the bright children who learn rapidly and earn extra promotions made high scores on the test, while the over-age plodders, in spite of the extra years they have spent in school, show their mental inferiority by making low scores. Of the 44 children who were seventeen years of age, 82 percent made scores below the median for the whole group, while of those twelve years of age, 82 percent were above the median.

TABLE IV. OTIS SCALE SCORE AND RATING IN WORK IN PERCENTS[a]

Percentile (Otis Scale)	School Work Rating[b]							Totals
	7	6	5	4	3	2	1	
95 to 100	0	0	2.0	18	26	30	24	100
85 to 95	0	0.9	1.9	18.8	43.5	30.6	3.9	99.6
65 to 85	0	1.4	12.3	40.3	28.0	15.2	2.4	99.6
35 to 65	0.6	3.2	14.1	46.0	23.0	11.5	1.3	99.1
15 to 35	2.9	11.3	28.5	36.4	15.7	3.4	0.9	99.1
5 to 15	6.9	22.7	29.7	21.7	17.8	0.9	0	99.7
0 to 5	7.8	37.2	35.2	17.6	1.9	0	0	99.7

[a] Read table as follows: of the highest 5 percent of the group according to Otis test scores, none were rated 7 in school work, none were rated 6, 2 percent were rated 5, etc.
[b] "1" in school work is very superior,
"2" in school work is superior,
"3" in school work is slightly above average, etc., see above.

TABLE V. OTIS GROUP INTELLIGENCE SCALE SCORE AND CHRONOLOGICAL AGE. 1,043 GRADUATING ELEMENTARY-SCHOOL CHILDREN[a]

Score	Ages							Totals
	11[b]	12	13	14	15	16	17	
150–159	.....	1	.....	.....	.....	.....	.....	1
140–149	.....	1	4	2	1	.....	.....	8
130–139	1	1	7	5	3	2	.....	19
120–129	.....	1	13	22	9	3	.....	48
110–119	.....	4	24	36	15	5	1	85
100–109	.....	1	19	44	34	8	.....	106
90–99	.....	3	23	54	56	14	4	154
80–89	.....	2	30	52	52	12	4	152
70–79	.....	1	15	53	57	28	5	159
60–69	.....	1	7	42	46	27	8	131
50–59	.....	.....	4	24	31	24	9	92
40–49	.....	.....	.....	11	24	13	5	53
30–39	.....	.....	.....	2	8	7	4	21
20–29	.....	.....	.....	4	2	4	2	12
10–19	.....	.....	.....	.....	.....	.....	2	2
Totals.....	1	16	146	351	338	147	44	1,043
Median Score....	134[c]	110	94	84	80	70	60	83

$r = -0.44$.

[a] Read table as follows: One child aged eleven scored between 130 and 139; one aged twelve scored between 150 and 159; one aged fifteen scored between 140 and 149; etc.

[b] Age eleven includes those from ten years and six months to eleven years and five months; age twelve from eleven years and six months to twelve years and five months.

[c] Score of one case.

Figure 2 compares the range of the middle 50 percent of each age group. It would be hard to present data more conclusive than these regarding the value of the test as a measure of the capacity for accomplishment in school work as it is now constituted.

The administrative uses to which intelligence-test results may be put are many. They may be used in studying pupils in groups; that is, the human material in one school can be compared with that in another. The classes of the same grade and of different grades within a school can also be accurately measured as to general capacity. The group intelligence test has already been used

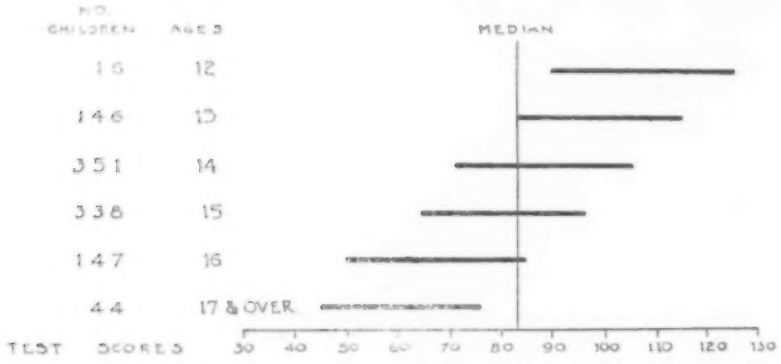

FIGURE 2. RANGE OF SCORES BY AGES
MIDDLE FIFTY PERCENT

successfully in some of the Oakland schools as an aid in making pupil segregations on the basis of ability to do school work.

The test results are also of value in studying individual pupils. The results are not as reliable in the measurement of individuals as in the measurement of classes, but if they are used cautiously, they are sufficiently accurate to render valuable information concerning individual pupils. Following are some of the ways in which test scores may be used in dealing with individual pupils:

1. A low score may help to confirm a teacher's judgment that a child's failure in school is due to inferior capacity. Special steps can then be taken to insure the best development of what ability the child is found to possess.

2. A high score associated with poor school work should call for special study. The child may be failing because of lack of application; or such factors as excessive diffidence, poor sight, defective hearing, or bad home conditions may explain why a child sometimes makes a high score and yet does poor school work. At any rate, all such cases are worth careful investigation to see if conditions cannot be remedied so that the child will work up to capacity. Cases are already on record in the Oakland schools where the attitude of a child toward school and the character of his school work have both been greatly improved as a result of a special case study made after a group test had been given.

3. The test scores often indicate that children should be considered for extra promotions. When a child makes an exception-

ally high score, and it is the judgment of his teachers that he is capable of doing work in a higher grade, he should be carefully studied with a view to advancement. If no good reasons seem to exist why he should not be put ahead, he should receive the extra promotion, due provision being made on account of it for the extra help he will need. On the other hand, when a child's test score and school work seem to agree, there is reason for satisfaction that the school is accomplishing what it should.

4. Test scores are valuable to teachers in educational and vocational counseling. It is only on the basis of exact knowledge that any effective advice can be given to children concerning their future educational and occupational careers. The intelligence test is one of several "lines" on a pupil's characteristics which a school counselor should have at hand when advising him.

5. The whole question of the differentiation of the curriculum in terms of the capacity and needs of pupils is one in which test results can lend valuable aid. This problem is already a pressing one in the high schools of some of our cities, for these schools are no longer receiving pupils selected almost entirely from the upper 25 percent of the general population. Indeed in some cities the freshman high-school class already has an enrollment 50 percent as large as that of the first elementary grade. Various cities report that practically all elementary-school graduates enter the high school.[1] In short, our high schools are now attended by many pupils with a mental capacity that is only average or even below average. At the same time, many high schools have little to offer these children except a curriculum designed for the highly selected college-preparatory group that attended high school a few decades ago. How miserably this curriculum fails, the elimination figures of most high schools clearly show. The problem of creating a new curriculum for children of average and below-average capacity is one on which intelligence tests give valuable aid.

These are but a few of the administrative uses to which group intelligence-test data may be put. Caution is, of course, necessary. Tests have no greater enemy than the inexperienced and enthusiastic individual who rushes into a program of testing and accepts all test results at their face value.

[1] Of the pupils included in this study, 87 percent entered high school.

THE VALUE OF THE WRITTEN PARAGRAPH SUMMARY

CHARLES E. GERMANE
College of Education, Des Moines University

A cursory glance through many of our leading textbooks on methodology shows that there is a wide difference of opinion among educational experts regarding the value of summarizing or outlining as a method of study. Some writers claim that asking students to summarize, outline, or make a synopsis of their lesson during the study period insures the following: (1) the pupils discriminate in their reading and (2) they organize what they have read.

On the other hand, many authors claim that much time is wasted in the study period by so much indiscriminate note-taking; that pupils are too much addicted to the "pencil and paper" habit; and that time could be more profitably spent by summarizing the lesson in class at the end of the recitation period.

However, since neither group of writers had any scientific evidence to offer, an attempt was made to determine experimentally the value of one form of summarizing as a method of study. The following is a brief synopsis of this experiment.

PROBLEM

Briefly stated the problem is: What is the relative value of a written paragraph summary of an article compared with the rereading of the same article for the same length of time.

METHOD OF EXPERIMENTATION

Selection of material.—Since this experiment was to be conducted in grades VI, VII, and VIII of the public schools it was necessary to select material that was neither too difficult for grade VI nor too easy for grade VIII. After considerable preliminary experimenting in the University Elementary School and the Junior High School of the State University of Iowa, the four following selections were chosen: (1) an article on the "Sweating System" adapted from Towne's *Social Problems*; (2) an article on "Immigration" adapted from Beard and Bagley's *American History;* (3) an article on "What We Should Know About Tuberculosis"

adapted from a bulletin issued by the New York State Department of Health; and (4) an article on the "Raising of Peanuts" adapted from the *Farmers' Bulletin, No. 431*, issued by the United States Department of Agriculture. These articles varied in length from seven to nine pages and were particularly suited to this type of problem in that they were all highly factual and logically organized.

Division of the class.—In order to measure the relative value of the written paragraph summary as compared with re-reading, it was necessary to divide the pupils into two groups of as nearly equal comprehension ability in silent reading as possible. It was desirable that the device for the division of the class should be like the device for measuring the comprehension of the pupils in the later tests. Hence it was thought that a division based on the ability of the pupils to read and comprehend one of the four articles previously mentioned would be as reliable a method as any. The article on the sweating system was selected as being best suited for this purpose, since very few of the pupils in these grades had ever read anything on that subject and many of them did not know what the title meant. The pupils were asked to read the article through twice and then were given a rigorous test consisting of questions of major and minor importance covering the entire selection. The scores made on this test formed the basis for the division of the pupils into the two groups.

Further to validate and justify this method of dividing the class into two groups the four articles named above were given to a group of twenty students under similar conditions. The instructions for the four tests were identical, the pupils being asked to read the article twice and then take the test. Table I shows the ranks of the pupils on the four different tests.

According to the Spearman footrule method of computation, the correlations are as follows:

Sweating System and Peanuts	0.42
Sweating System and Immigration	0.70
Sweating System and Tuberculosis	0.69

By using the table in Rugg's *Statistical Methods Applied to Education* to transfer these Spearman correlations into Pearson correlations, the following results are obtained:

Sweating System and Peanuts	0.64
Sweating System and Immigration	0.90
Sweating System and Tuberculosis	0.89

TABLE I. SHOWING THE RANKS OF THE TWENTY PUPILS ON THE FOUR TYPES OF READING MATERIAL

Sweating System	Peanuts	Immigration	Tuberculosis	
1	5	2	2	
2	13	5	5	
3	1	3	1	
4	3	1	4	
5	6	6	3	Q_4
6	2	7	8	
7	16	4	9	
8	4	10	7	
9	14	14	12	
10	10	9	14	
11	11	11	10	
12	15	15	17	
13	9	8	11	
14	8	13	13	
15	7	19	15	Q_1
16	18	12	6	
17	20	16	16	
18	19	18	18	
19	12	17	20	
20	17	20	19	

Read table thus: The pupil who ranked 1 in the test on the sweating system ranked 5 on peanuts, 2 on immigration and 2 on tuberculosis, etc.

These correlations, while not perfect, are high enough to indicate that in this group of pupils, ability to comprehend one article is closely correlated with ability to comprehend the other articles. However, the relationship can better be seen by comparing the quartile distributions for the four selections. While the same student does not maintain exactly the same rank in all four tests, yet it may be observed that there is very little change in the quartile distribution. Those students who are in the upper quartile for the first test are almost without exception to be found in the same quartile for the remaining three tests.

Method.—This experiment was conducted in grades VI, VII, and VIII in three representative public schools of cities in Iowa. The number of cases reported is 900. The Summarizing Group was called Group A, and the Re-reading Group was called Group B.

Both sections worked at the same time but in different rooms. Each spent thirty minutes in studying the material and at the end of that time each was subjected to the same rigid test of questions covering the content of the article. The tests were all graded by the writer and a graduate student who had worked through the experiment and who knew it thoroughly. The experimenters alternated the groups, one conducting the experiment in Group A one day and in Group B the next day, in an attempt to offset the effect of personality.

TABLE II. A COMPARISON OF THE TOTAL SCORES MADE BY THE SUMMARIZING GROUP (A) AND THE RE-READING GROUP (B) OF GRADES VI, VII AND VIII ON THE READING MATERIAL, PEANUTS

Grades and Groups	Whole Test	First Half	Reason Questions	Minor Questions
VI				
Group A	539	339	135	73
Group B	561	327	113	75
Group B excels by 4.1 %				
VII				
Group A	600	335	144	98
Group B	624	339	161	88
Group B excels by 4 %				
VIII				
Group A	752	407	243	85
Group B	799	462	230	94
Group B excels by 6.2%				

Read table thus: In grade VI, the aggregate score for Group A was 539 on the whole test, 339 when only the first half was considered, 135 on the reason questions, and 73 on the minor questions, etc.

In order that the pupils in the Summarizing Group (A) would understand exactly what was meant by summarizing, ten minutes were spent in class in summarizing two short paragraphs. The experimenter indicated the main points and emphasized the importance of speed and brevity. Then the pupils were given the article and instructions for the experiment. The pupils in

Group A were asked to read the article through *once*, rapidly and carefully. Then on the summarizing form provided, they wrote the main points or ideas of each paragraph. Thirty minutes were allowed for the reading and summarizing and at the end of that time the pupils were submitted to a rigorous fifteen-minute test.

The pupils in the Re-reading Group (B) were simply asked to read the article given them as many times as possible in the thirty-minute period. The same test was given them at the end of that time as was given to the pupils in Group A.

Results

A careful analysis was made of the written work of the pupils in Group A and it was found that in many cases the summaries were not completed in the time allotted. Especially was this

TABLE III. A COMPARISON OF THE TOTAL SCORES MADE BY THE SUMMARIZING GROUP (A) AND THE RE-READING GROUP (B) OF GRADES VI, VII, AND VIII, ON THE READING MATERIAL, IMMIGRATION

Grades and Groups	Whole Test	First Half	Main Questions	Minor Questions
VI				
Group A	371	300	25	35
Group B	379	284	31	29
Group B excels by 2.1%				
VII				
Group A	457	330	66	24
Group B	506	350	72	26
Group B excels by 10.7%				
VIII				
Group A	780	485	190	49
Group B	932	614	209	66
Group B excels by 19.5%				

Read table thus: Group A of grade VI made an aggregate score of 371 points when the whole test was considered, 300 points when only the first half was considered, 25 points when only the main questions were counted, and 35 points when only the minor questions were counted, etc.

TABLE IV. A COMPARISON OF THE TOTAL SCORES MADE BY THE SUMMARIZING GROUP (A) AND THE RE-READING GROUP (B) OF GRADES VI, VII, AND VIII ON THE READING MATERIAL, TUBERCULOSIS

Grades and Groups	Whole Test	Main Questions	Minor Questions
VI			
Group A	597	227	60
Group B	597	234	52
Group A and B are equal			
VII			
Group A	622	241	61
Group B	649	261	43
Group B excels by 4.3%			
VIII			
Group A	890	350	81
Group B	947	378	82
Group B excels by 6.4%			

Read table thus: Group A of grade VI made an aggregate score of 597 points when the entire test was considered, 227 points when only the main questions were considered, and 60 points when only the minor questions were considered, etc.

true in grades VI and VII. Therefore, because all completely summarized the first half of the material, a comparison was made of the scores of the two groups over the first half of the tests. It must be remembered, however, that such a comparison gives an advantage to Group A since many of the pupils in it spent the major portion of the thirty minutes on the first part of the article.

Again, some of the questions called for considerable thought and organization and consisted of four or five parts, while others were pure memory questions. An analysis was made of the answers of both groups to these two types of questions to see whether those who summarized answered more of the reason and main questions than those who re-read the article, or more of the memory questions.

The results for the two groups on the three types of reading material are summarized in Tables II, III, and IV.

The Delayed Recall

It was thought that probably the pupils in the group which summarized would retain what they had read better than those

in the group which merely re-read the articles. To see whether or not this was so, the same test as was given in the first place was given to both groups at the end of an eight-week period. The results are summarized in the following tables.

TABLE V. A COMPARISON OF TOTAL SCORES OF THE SUMMARIZING GROUP (A) AND THE RE-READING GROUP (B) OF GRADES VI, VII, AND VIII ON MATERIAL, PEANUTS

Group	Grade VI	Grade VII	Grade VIII
A	201	216	373
B	212	218	390

Read table thus: Group A in grade VI made an aggregate score of 201 points on the test on peanuts, grade VII, an aggregate of 216 points, etc.

TABLE VI. A COMPARISON OF TOTAL SCORES OF THE SUMMARIZING GROUP (A) AND THE RE-READING GROUP (B) OF GRADES VI, VII, AND VIII ON MATERIAL, IMMIGRATION

Group	Grade VI	Grade VII	Grade VIII
A	92	148	354
B	105	139	382

Read table thus: Group A in grade VI made an aggregate score of 92 points on the test on Immigration, Group A in grade VII, an aggregate of 148 points, etc.

A careful comparison of the scores of Groups A and B as presented in Tables V and VI shows that Group B excels in five cases. Group A in grade VII excels on the tests on Immigration.

Conclusions

1. The data presented in the preceding tables would seem to indicate:

   a. That the re-reading of an article is a more economical method of study than the written paragraph summary, when each group studies the same amount of time.

   b. The above conclusion is further substantiated by the fact that the Re-reading Group almost invariably made higher scores on the immediate recall test than the Summarizing Group whether the results were compared over the entire test, the first half of the test, over those questions called "reason" and "main" questions, or over those styled "minor" questions.

2. A careful examination of the pupils' summaries showed:

a. That verbosity in expressing the main points characterized much of the work of grades VI and VII, and hence much time was wasted.

b. That much time was wasted in grades VI and VII because of the inability of the pupils to differentiate between essential and non-essential points.

3. The Re-reading Group in grade VIII made higher scores in every case than the Summarizing Group even though the summaries of the pupils in the Summarizing Group were invariably completed and of an excellent quality.

4. Group A in grade VIII could easily have excelled had the pupils been able to recall at the time of the test the main points found on their summarizing forms. Hence it would seem that the Summarizing Group was handicapped by the method of summarizing each paragraph since the pupils seemed to have a large number of isolated facts but failed to see the relationship and grasp the article in its entirety.

5. Time would be very profitably spent in emphasizing comprehension in the reading of all school material. This statement is verified by an analysis of the papers of the pupils. While the students read the articles on an average of three or four times, yet the median score was less than one-third of the possible score.

6. The data presented in Tables V and VI showing the results of the delayed recall tests would seem to indicate:

a. That since all retained only a comparatively small number of the ideas in the original material, frequent reviews would be advisable as a part of the school routine.

b. That the written paragraph summary does not seem to be as efficient a method as the re-reading of an article.

7. Observation of the Re-reading Group at work showed that after the pupils had read the article a couple of times, they apparently lost interest in it. It is probable that the pupils of this group could have read a longer article—one which would have held their interest during the entire period—just as profitably in the same length of time.

AGE AND GRADE NORMS FOR THE NATIONAL INTELLIGENCE TESTS, SCALES A AND B

LEWIS M. TERMAN AND ETHEL D. WHITMIRE
Leland Stanford Junior University

In June, 1920, the National Intelligence Tests, Scales A and B, Form 1, were given to all the pupils between the ages of eight and fifteen inclusive, who were present on certain days in the public schools of Vallejo, California. They were given by Dr. Terman, Superintendent A. C. Barker, and four members of the school force who had had experience in giving group tests. The scoring and tabulating were done by university students and high-school seniors who were paid for their time. Their work was carefully supervised by Miss Whitmire, but sufficient funds were not available to have the papers scored twice. All computations were carefully checked.

In all cases Scale A was given before Scale B. In about half the classes of grades V to VIII both scales were given at a single sitting, broken only by a rest of five to ten minutes. This was recognized as undesirable but was unavoidable because the tests were made during the last week of the school year. In the other classes Scale B followed one day after Scale A.

Giving both scales at a single sitting does not, of course, affect the norms thus secured for the scale given first (Scale A), nor do we believe that it affected appreciably the norms for Scale B. It could only do so by making the norms too high (effect of practice) or too low (the effect of fatigue). In this case, since the tests of Scale B differ in kind considerably from those of Scale A, fatigue effects would ordinarily be expected to outweigh the practice effects. Whether they did so to any considerable extent can be answered by comparing for each grade and each age separately the ratio of the Scale B norm to the Scale A norm in our own data with the corresponding ratios secured at Washington and Pittsburgh, where in all cases Scale B followed Scale A by one day. These ratios are shown in Tables I and II. It will be seen that on the

[1] The "Age and Grade Standards (Norms)" as printed in the Manual of Directions accompanying the National Intelligence Tests were "based upon approximately 2,000 examinations made in Washington, D. C., and 2,000 made in Pittsburgh, Pennsylvania."

whole the ratios for the various ages (Table I) differ but little from those of Washington and Pittsburgh,[1] and that in those ages where the two eastern cities differ most from each other, Vallejo occupies an intermediate position. The ratio differences by grade (Table II) are more irregular and comparison is rendered difficult by the fact that the Washington norms are given by half grades. However, there is little if any evidence that our norms have been affected by the fact that both scales were sometimes given at the same sitting.

TABLE I. RATIO OF SCALE B SCORES TO SCALE A SCORES, BY AGES[a]

Age	Washington	Pittsburgh	Vallejo
8	106.16	94.32	102.9
9	102.74	95.61	101.2
10	103.66	97.92	101.3
11	100.00	99.08	100.7
12	98.23	100.00	98.2
13	100.00	100.78	101.1
14	95.45	100.00	100.5
15	97.54	98.38	97.7

[a] Ratios are multiplied by 100. Those greater than 100 indicate that the Scale B score was higher than the Scale A score; those less than 100 indicate the reverse.

TABLE II. RATIO OF SCALE B SCORES TO SCALE A SCORES, BY GRADES[a]

Grade	Washington	Pittsburgh	Vallejo
IIIA	.....	.....	94.7
IIIB	110.34	.....	.....
IVA	104.76	96.48	103.5
IVB	105.06		
VA	129.85	96.85	101.5
VB	101.03		
VIA	100.00	100.00	99.4
VIB	99.14		
VIIA	100.85	97.64	96.2
VIIB	98.47		
VIIIA	97.10	100.71	98.9
VIIIB	96.48		

[a] Ratios are multiplied by 100. Those greater than 100 indicate that the Scale B score was higher than the Scale A score; those less than 100 indicate the reverse.

Inasmuch as the primary purpose in giving the tests was to secure data for tentative age norms, effort was made to test every pupil who had reached the eighth birthday and had not yet become sixteen. In order to accomplish this all the pupils in grades III to VIII were tested. Tests were then given to all pupils in the high school who had not reached the age of sixteen, and to all pupils in the first two grades who were eight years old or older.

According to the state compulsory school law in force at the time the tests were given, every pupil in the state who had reached the age of eight years and who was not yet sixteen should have been in school. Reference to Table III, however, will show that some selection has taken place by the age of fourteen, and that by the age of fifteen it is considerable. To what extent this is due to partial failure of the compulsory school law and to what extent it is due to the attendance of pupils in private and parochial schools, it is impossible to say; nor do we know the nature of the selection that has taken place. However, our norms for ages eight to fourteen inclusive are probably fairly satisfactory for the city in question. Those for years eight and nine are certainly far more valid than those published in the Manual of Directions for Pittsburgh and Washington, as the latter included no tests below the upper half of the third grade. Children of these ages who are in the upper half of the third grade are probably bright children. Accordingly, the Washington and Pittsburgh scores for eight- and nine-year-old children are too high. On the other hand, for ages above fourteen the Pittsburgh and Washington norms are quite worthless, since no tests were made above the eighth grade. Thus the brighter children were left out and the norms for these ages are too low. Even the norms for fourteen from those cities are of questionable value, for in the case of Washington the number of pupils at this age is only 60 percent and in Pittsburgh only 56 percent of the average number for the three previous years. The corresponding figure for Vallejo is 79 percent. Even year fifteen in Vallejo shows a retention of 67 percent as compared with the average for the ages 11, 12, and 13.

As regards social and economic status Vallejo is probably not far from the average of California cities of similar population, though it should be remembered that the general average for California is high as compared with the majority of other states. The only important industries of the city are United States navy

TABLE III. NUMBER OF CASES INCLUDED IN AGE AND GRADE AVERAGES

Grade	Washington	Pittsburgh	Vallejo	Age	Washington	Pittsburgh	Vallejo
III	159	.....	175	8	84	14	171
IV	308	211	210	9	177	96	200
V	291	221	228	10	237	189	172
VI	316	233	142	11	276	213	185
VII	264	182	183	12	280	169	184
VIII	209	196	135	13	252	182	159
IX	.....	.....	102	14	164	105	144
X	.....	.....	38	15	52	41	119
Total	1,547	1,043	1,213	Total	1,522	1,009	1,334

yard activities (largely shipbuilding) and flour milling. These industries have given the city a large proportion of semi-skilled and skilled labor, chiefly of American stock. There is no reason to believe that the proportion either of unskilled laborers or of the professional and semi-professional classes is abnormally large.

The following figures show the nationality classification of the 18.65 percent of school children whose parents were both foreign born:

Extraction	Percent
Latin	6.71
Teutonic	4.73
Scandinavian	3.57
Slavic	2.26
Miscellaneous	1.38
Total	18.65

Children of foreign-born parents were included in the calculation of norms, as nearly all such children were born in the United States.

Tables IV and V make possible a comparison of the age and grade norms of Scale A, Scale B, and total of A and B for the three cities. In all cases these norms are based on averages. For the convenience of those who prefer to use medians the age and grade

medians for Vallejo are given in Table VIII. In Table IV it will be observed that our average total scores for grades IV, V, and VI correspond very closely to those from Washington, but that grades VII and VIII are each about 12 points lower than the corresponding grades in Washington. A difference of 12 points corresponds roughly to about 0.4 of a grade, the average difference from grade to grade being approximately 30 points.

It should be noted that the tests on which the grade averages were based were, in all three cities, made near the close of the school year.

TABLE IV. GRADE AVERAGES FOR WASHINGTON, PITTSBURGH, AND VALLEJO

Grade	Washington			Pittsburgh			Vallejo		
	Scale A	Scale B	Total	Scale A	Scale B	Total	Scale A	Scale B	Total
IIIA	.....	.....	.....	.....	.....	.....	51.2	48.5	100
IIIB	58	64	122	.....	.....	.....			
IVA	63	66	129	85	82	167	71.9	74.4	146
IVB	79	83	162						
VA	67	87	174	95	92	187	89.4	90.7	180
VB	97	98	195						
VIA	104	104	208	112	112	224	107.9	107.2	215
VIB	116	115	231						
VIIA	118	119	237	127	124	251	120.8	116.0	237
VIIB	131	129	260						
VIIIA	138	134	272	140	141	281	132.7	131.3	264
VIIIB	142	137	279						

No comparison of age averages need be made except for the years 10 to 13, as the Washington and Pittsburgh norms are misleading and worthless below 10 and above 13. For the years 10 to 13 the Vallejo norms are below those of Washington by 15, 16, 11, and 4 points, respectively, and for the same four years they are below those of Pittsburgh by 38, 35, 13, and 23 points. That is, the Vallejo norms for these four ages average nearly a year below those of Pittsburgh.

TABLE V. AGE AVERAGES FOR WASHINGTON, PITTSBURGH, AND VALLEJO

Age	Washington			Pittsburgh			Vallejo		
	Scale A	Scale B	Total	Scale A	Scale B	Total	Scale A	Scale B	Total
8	65	69	134	88	83	171	34.6	35.6	70
9	73	75	148	91	87	178	57.1	57.8	115
10	82	85	167	96	94	190	75.7	76.7	152
11	98	98	196	108	107	215	89.6	90.2	180
12	113	111	224	113	113	226	107.4	105.5	213
13	119	119	238	128	129	257	116.3	117.6	234
14	132	126	258	130	130	260	133.3	133.9	267
15	122	119	241	123	121	244	137.0	133.8	271

TABLE VI. GRADE INCREMENTS IN TOTAL SCORES OF SCALES A AND B

Grades	Washington	Pittsburgh	Vallejo
IV to V	36	20	34
V to VI	36	37	35
VI to VII	29	27	22
VII to VIII	17	30	27
Average	29	28.5	29.5

TABLE VII. AGE INCREMENTS IN TOTAL SCORES OF SCALES A AND B

Ages	Washington	Pittsburgh	Vallejo
8 to 9	+14	+ 8	+45
9 to 10	+19	+12	+37
10 to 11	+29	+25	+28
11 to 12	+28	+11	+33
12 to 13	+14	+31	+21
13 to 14	+20	+ 3	+33
14 to 15	−17	−16	+ 4
Average for 9 to 14	22	16.65	30.65

Tables VI and VII show the age and grade increments in total score of Scales A and B in the three cities. Since the Washington norms are given by half grades, the grade increments for that city are based upon the averages for the upper half of each grade.

The grade increments for all three cities suggest that a year of growth ought to yield an increment of about 29 points. However, when we examine the age increments of Table VII we find that only those of Vallejo give results which agree with this assumption. Washington and Pittsburgh show increments at most ages which are considerably less than this. In both of these cities the increments are small in the lower years, increase gradually until toward the middle of the age range they approximate the expected increments, then decrease in the higher ages until finally the increment becomes really a decrement corresponding to about a half year of mental age. This is, of course, due to selection. It is evident that at Washington and Pittsburgh only the brighter children below ten years were tested. On the other hand, at Washington only the duller children above twelve and at Pittsburgh only the duller children above thirteen were tested. At Vallejo, however, the annual increments from ages nine to fourteen approximate those which would be expected according to the grade results. Here the low increment from fourteen to fifteen is probably due in the main to the selection which has operated with our fifteen-year-olds. The unexpectedly large increment at Vallejo from eight to nine is probably due to the fact that many eight-year-olds are not adequately measured by the scale, perhaps because of its literacy requirements. This would tend to reduce the average for year eight considerably below where it ought to be, and is an error which the median tends to exaggerate even more than does the average. At any rate our comparison of the age and grade increments suggests that our age norms for the years nine to fourteen are probably not far wrong for the city in question, and that those of Washington and Pittsburgh are probably of little value except in the very middle of the age range.

The author of the Manual of Directions has made a set of provisional age norms in which an effort has been made to allow for the factor of selection in the upper and lower ages. The method of making these corrections is not stated, but the corrections themselves are shown in the third column of Table VIII. In the fourth and fifth columns of the same table are shown the age

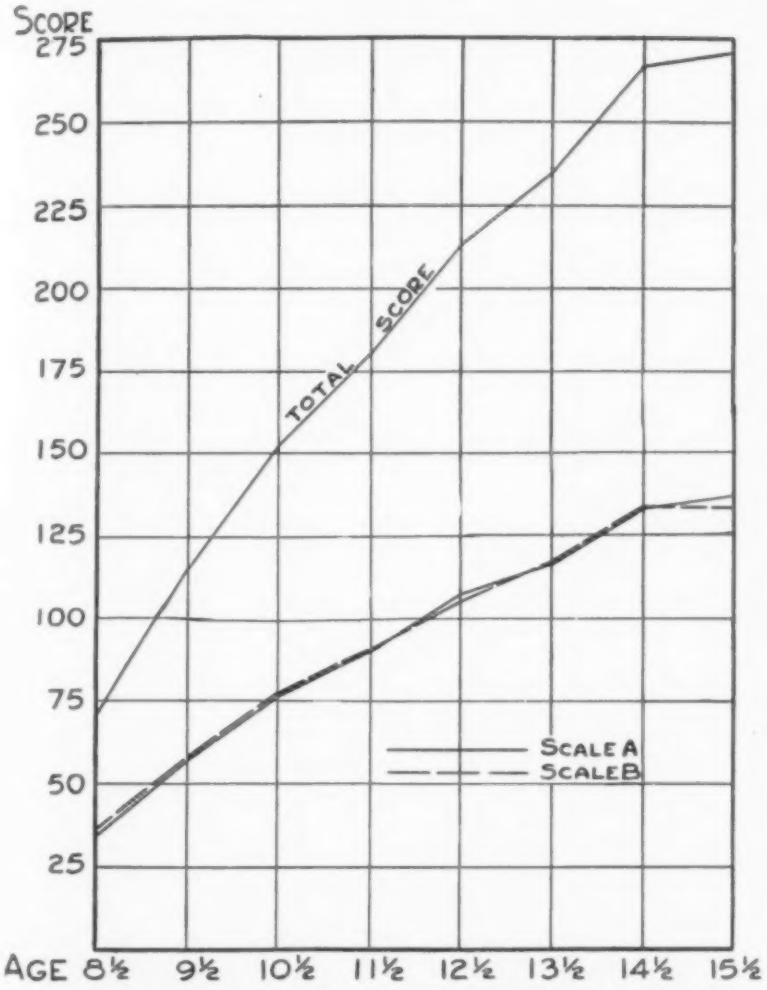

FIGURE 1 Scale A, Scale B, and Total of A and B by age. Children from 8 to 9 are classed as 8½ etc.

averages and age medians for Vallejo. Until extensive data are available on a more nearly unselected group, we believe that the age norms given in columns 4 and 5 should be used in preference to the figures in the third column, which are the tentative norms recommended in the Manual of Directions.

TABLE VIII. AGE NORMS FOR TOTAL SCORE OF SCALES A AND B

Age[a]	Washington Average	Pittsburgh Average	Norms Recommended in Manual	Vallejo Average	Vallejo Median	Average for Scale A Alone
	1	*2*	*3*	*4*	*5*	*6*
8.5	134	171	130	**70**	**60**	**35**
9.5	148	178	155	**115**	**118**	**57**
10.5	167	190	180	**152**	**158**	**76**
11.5	196	215	205	**180**	**186**	**90**
12.5	224	226	225	**213**	**216**	**107**
13.5	238	257	245	**234**	**243**	**116**
14.5	258	260	260	**267**	**268**	**133**
15.5	241	244	275	**271**	**274**	**137**

[a] In all cases age 8 includes children from 8 to 9, age 9 those from 9 to 10, etc. Accordingly, the age standards given are always those for ages 8.5, 9.5, etc. This more correct notation has been used in Table VIII and Figure 1.

The age medians for Scale A and Scale B are very nearly equal at all ages when Scale A is given first. If a pupil has taken only Scale A, and it is desired to estimate what his probable total score would have been had he taken both scales, it is only necessary to multiply the score on Scale A by 2. While this rule is not absolutely accurate, it is sufficiently so for all practical purposes. The correlation between Scale A and Scale B was computed and found to be 0.928, Pearson.

Figure 1 shows graphically the Vallejo age curves for Scale A and Scale B. These curves are based upon averages. Within the range of validity (probably 9 to 14 years) the curve of each scale approximates a straight line. If it were not for the reduced validity of the test below 9 years and the unfavorable selection of cases above 14, it is probable that the line would be approximately straight from 8 to 15.

THE VALUE OF A HANDWRITING SCALE TO AN UNTRAINED TEACHER

R. L. MORTON
Ohio University

Various investigators have shown that the use of either the Ayres or Thorndike handwriting scale will materially reduce the variation among judges in marks assigned to the same specimens, and will also reduce the variation between successive markings made by the same individual. This fact of reduced variability would imply increased accuracy. A few teachers, however, are consistent in their ratings of handwriting but grossly inaccurate—that is, they consistently mark 10, 15, or 20 percent above or below the true value. This consistent inaccuracy would not be particularly serious if it existed to the same degree and in the same direction among all teachers, or even among all teachers in the same grade in the same school system. This, however, is not the case.

The writer has, during the past two years, gathered data bearing on this question: Will the use of a handwriting scale, among those who have never used it before, result in more accurate grading of handwriting? In the school years 1918–1919 and 1919–1920, students in the writer's classes in educational measurements worked out an experiment which is an adaptation of that suggested by Thorndike in the *Teachers College Record* for November, 1914.

Students first rated Thorndike's fifty specimens (whose true values are known) by the usual percentage method. They were told simply to rate these specimens as they would, or did, rate handwriting in their schools; except for the fact that they were not to think of the writing as belonging to any particular grade but were to think of 100 percent handwriting as that which is perfect and 0 percent handwriting as that which is worthless, or possessed of no beauty or legibility. The students were then supplied with copies of the scale. For over 90 percent of them it was the first time that they had ever seen a scale, while only four or five of the entire number had ever actually *used* a scale before. A brief explanation of the scale and its uses was given by the instructor and the students proceeded to rate the fifty specimens

again, this time in terms of scale units. These grades were set down in appropriate columns of a table which each student constructed, and the instructor then supplied the correct values of the fifty specimens as given by Thorndike, which were also recorded.

Scale ratings were then converted into percents by multiplying by 5.5 in the case of the Thorndike scale, and in the case of the Ayres scale by first adding 20 to the scale values and then multiplying by 0.825. The deviation of each score from the correct score was then computed; these deviations were added; and the sum was divided by 50. The resultant quotient represented each individual's average error in scoring the same specimens with and without the scale. To facilitate computation, average errors were expressed to the nearest integer.

During the year 1918–1919, 176 students performed this experiment using the Thorndike scale; and during 1919–1920, 102 students did so using the "Gettysburg Edition" of the Ayres scale. As will be described later 44 others performed the experiment under somewhat altered conditions. All these 322 were students in the writer's classes in educational measurements conducted in sixteen cities and towns in southeastern Ohio in connection with the work of the Extension Department at Ohio University. Most of them were teachers in active service in the grades. A few were high-school teachers, principals, and superintendents. A total of 31,100 ratings were used.

Table I shows the distribution of the average errors of 278 of these teachers in both directions. Table I is to be read as follows: One teacher made an average error of 7 percent without the scale and 4 percent with it, one an average error of 8 percent without a scale and 4 percent with it, three an average error of 9 percent without a scale and 4 percent with it, and so forth. It will be seen that the median of average errors without a scale is 11.8 percent and with a scale the median is reduced to 6.8 percent. In other words, the mere use of a scale among these 278 teachers, who were almost entirely unacquainted with a scale and its use, reduced the average error 5 percent. Thus, 42.4 percent of the error was eliminated when the scale was first brought into use. Note that the lowest average error without a scale was 6 percent. Eight people, or 2.9 percent of the entire group, graded with this degree of accuracy by ordinary methods. On the other hand, 121, or 43.5 percent,

showed this or a greater degree of accuracy when a scale was supplied; and this was in spite of the fact that the scale was to them a new tool, to the use of which they were not accustomed.

TABLE I. DISTRIBUTION OF THE AVERAGE ERRORS (IN PERCENTS) OF 278 TEACHERS WITH AND WITHOUT A SCALE

With a Scale: Average Error	Without a Scale: 6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	Over 20	Total
4		1	1	3						1		1					7
5	4	2	3	6	5	3	4	1	4	3	1		2	1			39
6	3	4	8	11	6	5	7	5	5	7	5	1	3		2	3	75
7	1	3	8	9	4	4	6	7	1	7	2	3	1	2	1	4	63
8		1	3	7	5	4	4	2	4	3	3	1		2	2	1	42
9				1	2	1	3	3	2	1	2	1	1		1	2	20
10		1	3	2	1	1	3	1		2							14
11					2		2	1		2		1		1		2	11
12					1		1				1						3
13															1		1
14			1					1	1								3
Total	8	12	27	39	26	18	30	21	17	26	14	8	7	6	7	12	278

Median error without scale, 11.8 percent
Median error with scale, 6.8 percent

Of the 278 teachers, 13 or 4.7 percent showed a loss in accuracy when using a scale for the first time; 13 or 4.7 percent showed no change; and 252 or 90.6 percent showed an improvement, ranging from 1 percent to 28 percent. This condition is more clearly stated in Table II.

TABLE II. GAINS IN ACCURACY

Gain (Percent)	Frequency	Gain (Percent)	Frequency
−6	1	8	18
−5	0	9	15
−4	0	10	12
−3	1	11	7
−2	4	12	7
−1	7	13	6
0	13	14	3
1	28	15	3
2	29	16	3
3	27	17	0
4	27	18	1
5	25	19	1
6	23	28	1
7	16		
		Total	278

Our data furnish some evidence on the much discussed question as to which of the two scales is the more valuable. The 44 teachers (all in one city) referred to in a preceding paragraph scored only 25 of the 50 specimens but scored these first in the percentage fashion, then by the Thorndike scale, and finally by the Ayres scale. The brief period of practice with the Thorndike scale may have resulted in increased skill in scale use and hence have given the teachers some advantage in using the Ayres scale. On the other hand, since the work was all done at one sitting, fatigue probably offset this advantage. At any rate, the median of the 44 average errors with the Thorndike scale was found to be 6.4 percent and with the Ayres scale, 6.3 percent. These scores are not included in the tables.

As stated above, 176 of the 278 teachers listed in Table I used the Thorndike scale and 102 used the Ayres scale, "Gettysburg Edition." The median of the average errors made by the first group without a scale was 11.9 percent and with a scale 6.5 percent. For the second group the corresponding figures were 11.7 percent and 7.3 percent. These figures seem to indicate that untrained teachers can do slightly more accurate work with the scale devised by Thorndike.

Does experience in teaching make for increased accuracy in handwriting grading either with or without a scale? To answer this question I correlated the average errors without a scale of the 176 students in the 1918–1919 group with the number of years of teaching experience, using the Pearson formula. I found the value of r to be 0.0009. Teaching experience neither increased nor decreased on the average, the accuracy of handwriting grading without a scale, i. e., in the percentage manner.

To measure the effect of experience upon ability to use a scale without practice, experience in years was correlated with average errors with a scale. A positive coefficient of 0.24 was found. The reader should bear in mind that a positive coefficient here means that, so far as this group is typical, the more experience a teacher had, the less profitably she used a handwriting scale without practice. In other words, the evidence is that teaching experience does not increase the average individual's accuracy in grading handwriting by the ordinary percentage method under the conditions involved in the experiment, and that it actually decreases his ability to use a scale.

Editorials

THE SCHOOL AS A SELECTIVE AGENCY

Average mentality expressed in the form of an intelligence quotient is 100. There is evidence to show that the typical first-year high-school pupil has a mentality of 105 and that, three or four years later, high-school graduates show a record of 111. Between graduation from high school and entrance to college there is further evidence that mentality moves up something like four notches to 115. If the same process continues in college, the median intelligence quotient of college graduates will be found to be 120 or over.

The school has not engendered this ability. Those who evince a given degree of mentality have it as a possession as inherently as they have blue eyes or Roman noses. Unconsciously the school, year after year, sifts through meshes of larger and larger size those who are relatively less capable, and retains those whose endowments mark them for intellectual success.

As an institution devoted to preparing the young for more complete living, the school has always been so much an instrument of selection that it is a question whether those who receive its benefits are as much *developed* as they are *certified* for the affairs of life. The school not only trains children and young people but also attests their success as students. Its official record is accepted as evidence of competence. If those who have attended school a relatively long time have greater earning capacities than those who have attended a shorter time, if social standing is powerfully influenced by educational advantages, if appointment officials and civil service commissions give much weight to diplomas and degrees, and if in general we proceed on the assumption that each higher round of the educational ladder confers upon those who reach it additional fitness for life, it is because those who have more schooling are not only better trained but also more highly selected.

The notion that the school produces its results through training alone is too naïve to be entertained. It is true that we sometimes hear teachers and other school officers speak as if the suc-

cess of their pupils in after life were wholly attributable to instruction. Those who have sought to measure the results of teaching have often been told to look for them ten, twenty, or thirty years later. The fact is that our students have certain endowments long before they come under our influence. We may help them to realize their possibilities, or we may hinder them in their development; but we cannot create capacity.

Within the past few years the means have multiplied by which the school has become consciously the directive and selective agency which it has always been unconsciously. Our educational and intelligence tests permit us to ascertain the capacities of pupils far more accurately than ever before. Thus, the teacher becomes a guide and director. He is still the trainer of youth but he selects one to be trained in this way and another in that. He is no longer blind. He sees the limitations and the failures of his pupils and he applies his remedies—that is, he brings training into play—where it will do the most good.

Some will say that these conceptions are fatalistic and depressing, but the real question is not a question of fatalism but of truth. As to the teacher, it seems to us that his calling is likely to take on added importance because of the new opportunities which these conceptions will bring him. Instead of prescribing the same treatment for all, he will become the expert diagnostician. On the basis of mental ability he will reclassify children, and because of their special abilities, he will further subdivide them. Promotion will be for those who can profit by advancement rather than for those who have reached a predetermined standard. The gifted, the mediocre, and the dull will be detected, not merely for the sake of identifying them, but essentially in order that finer adjustments may be made and larger opportunities realized. When to the old consecration by which the craft has been distinguished is added the new insight which modern methods can confer, teaching will acquire new dignity. The modern school as an institution and the teacher as its official will then be recognized as properly concerned with selecting as well as with training.

B. R. B.

DERIVING AND PUBLISHING EVALUATED MATERIAL

The derivation of tests involves the evaluation of the elements of which the tests are composed. We need a far larger body of

these evaluated elements than now exists. Indeed, there is ground for believing that if all the elements which are now used in the teaching process were evaluated, a much more intelligent use would be made of them. For example, if all the problems in an arithmetic had a difficulty rating, it is clear that fewer misunderstandings between teachers and pupils would arise, that more reasonable assignments would be made, and that a more satisfactory gradation and adaptation of material to the needs of pupils would be possible.

And even if we confine ourselves to the use of test material for measuring purposes, it is perfectly evident that we need a far richer supply of these fundamental data. This is especially apparent when we consider the desirability of refining the existing tests. For example, at the meeting of the education section of the American Association for the Advancement of Science at Chicago on December 30, Dr. Luella W. Pressey presented an analysis of Monroe's Silent Reading Tests. Because of the relatively large number of exercises in the several parts and forms of this instrument, she was able to devise several tests, each made up of material of a similar character—one of science material, another of poetry, etc. She reached certain interesting conclusions as to the limitations of a general reading score based on a test consisting of different sorts of material. It was clear, however, that Dr. Pressey could have carried her investigation much further if she could have had at her command a still richer body of evaluated exercises. It has often been observed that one of the tendencies in the test movement is precisely towards this type of analysis. In virtue of it we shall, no doubt, have in the place of general reading scales, scales for narrative, for exposition, for scientific writing, for poetry, etc. We already have, instead of general scales for composition and handwriting, analytical scales involving the measurement of special abilities.

Not only in the refinement of our instruments and in the analysis of the more general forms of educational measurement, but also in the repetition of tests do we require a greater volume of material. Not infrequently a teacher or a superintendent, after having used all the available forms of a test, desires to continue its use in order to secure progress records. Since the children are already familiar with the existing material, their subsequent scores

will represent not only progress but also the learning effect of the previous presentations. Since the latter is difficult or impossible to measure, the progress record remains undetermined. And yet such a record constitutes an especially valuable result of testing. Probably one of the most serious criticisms of the use which school people are now making of tests is that they do not continue their use long enough to receive the full benefit of them. Yet in justice to the school people, it should be said that a more continuous use of the tests would doubtless be made if they were extended, enriched, and provided with more alternative forms—in other words, if a larger body of evaluated material were available.

Moreover, every research bureau should have at hand a supply of standardized words, problems, questions, and other items which it can utilize in new combinations and for special purposes. Indeed, it seems to us to be one of the routine duties of such a bureau annually or semiannually to gather the responses of a large number of children to a variety of test elements. Such material in the files of the bureaus should be as much a part of their standard equipment as typewriters or adding machines.

We shall not attempt to be exhaustive in pointing out the uses that may be made of scaled data. Our purpose is served if we establish the fact that these uses are many and important. The lack of such material is evident on every hand. Questions of import cannot be answered without it. The "best" method of teaching a given subject, the value of particular kinds of drill, the amount of time that may be most economically spent on a subject or a phase of a subject, the distribution of this time into long or short periods, the rate at which material should be presented, the influence of types of imagery on methods of learning and of teaching, the effect of re-learning or review and of the intervals between periods of re-learning, the comparative efficiency of a "whole method" and a "part method" of learning, the extent to which there is a transfer from abilities acquired in given situations to abilities manifested in other situations—these and almost countless numbers of problems which beset the paths of both school people and research workers demand for their investigation a rich and highly organized body of calibrated items.

Such being the case, it is proper to inquire how this sort of material can be produced in larger volume. Everyone who has at-

tempted the evaluation of data knows that the process is tedious and expensive—often prohibitively so when sufficient returns are to be handled to secure a relatively high degree of reliability.

But it is quite possible to abridge the usual method of evaluation and to ascertain the difficulty of new material by utilizing material whose difficulty has already been determined. We do not mean that the results obtained in this way will be as accurate as those obtained from a large number of returns on each item, but sufficient accuracy may be secured for most practical purposes. Thus, as money makes money, so evaluated material may be used to make evaluated material.

We have tried this sort of thing in a small way. On a list of fifty words we had returns from between 4,000 and 5,000 children in the fourth, fifth, and sixth grades.

We divided the list into two equal parts, assuming that the first twenty-five words were of known difficulty and the second twenty-five of unknown difficulty. We then investigated the returns from a small city system in which there were about one hundred children in each of the fourth, fifth, and sixth grades. We found that on the first twenty-five words these children did better than our norms led us to expect—better by an amount which we were able to determine. Now, to the extent that the spelling ability of this group of children was general and not merely specific with reference to these words, the same superiority might be expected to be revealed in the second half of the list—i. e., in the undetermined words. Accordingly, the results from this small city on the undetermined words were reduced by the same amount. Up to this point we had been ignoring the fact that we really knew the difficulty of the second half of the list. We now compared the difficulty as obtained from the three hundred children with the difficulty which we had previously found. Only two or three of the words were displaced as much as two columns on the Ayres scale, and most of them were not displaced at all. This result was obtained from about one-fifteenth of the original number of cases and with perhaps one-twentieth of the labor required to set up the first norms.

A purist may be pained at even this amount of divergence. The differences, however, are surprisingly small. In the first place the Ayres scale has a certain unreliability. Indeed, we have reason to believe that this unreliability is greater than has

been supposed. In the second place, our original determinations of difficulty for these fifty words were also unreliable to a certain extent. In the third place, scarcely any of the discrepancies amounted to more than half the average difference in spelling ability between successive grades. We should not recommend the use of evaluations of this sort for the more refined purposes of measurement, but for most practical purposes we believe them to be sufficiently accurate. If such is the case, it lies within the resources of almost any superintendent to evaluate his own material, if he has at hand a small amount of accurately scaled elements.

We believe, therefore, that all who are interested in educational research should be busy in the standardizing of material. Moreover, this material once standardized should not lie in the files of research bureaus nor in the desks of school superintendents. It should be published. Its production ought to be a cooperative enterprise. In order that it may be, publication is indispensable. Moreover, the material must be issued in such a form that others may add to it. Original rather than derived figures are necessary. For example, it is not sufficient to print merely percents correct for words or questions. The number of children participating and the number of correct responses must also be given, so that each of these figures may be augmented by other workers, and to the end that the derived percents may become progressively more reliable. But this sort of publication is expensive. It tends to run to long and very detailed tables. Such tables, however, are to be regarded as source material. Like the population tables and financial statistics of the Federal Census Bureau, they are fundamental to the study of a vast number of questions. Their independent value is not nearly as great as their value when brought into relation with other things.

A reasonable amount of this sort of material should be presented by the periodicals. The JOURNAL OF EDUCATIONAL RESEARCH proposes to do its part.

B. R. B.

Reviews and Abstracts

E. H. CAMERON, *Editor*

GUGLE, MARIE. *Modern junior mathematics*. New York: Gregg Publishing Company, 1920. 3 books.

The junior high school as a part of the 6-3-3 plan of school organization is winning strong approval and support from many quarters. Miss Gugle's attempt to provide this new type of high school with suitable modern texts in mathematics is one of the more successful among several such recent ventures. Book One, planned for the seventh school year, covers rapid correct calculation, checks and short cuts, and everyday business arithmetic with the elements of bookkeeping. Book Two, an eighth-year text, is given up to observational geometry with an introduction to algebra through the formulas arising from certain measurement problems. It contains an appendix on mathematics clubs. Book Three, for the ninth year, develops algebra through simple quadratics in two unknowns; applies sines, cosines, and tangents to problems in heights and distances; and closes with a short introduction to demonstrative geometry.

Some noticeable characteristics of these books are: (1) drill in rapid accurate calculation in everyday business arithmetic motivated through time tests; (2) continued and well-coordinated use of graphic methods; (3) a simplified treatment of percentage; (4) a close correlation between arithmetic and familiar everyday business affairs; (5) a concrete informal and interesting observational study of form and measurement which makes an excellent foundation for later work in formal demonstrative geometry as well as leading to formulas introducing in a natural way the literal notation of algebra; (6) the emphasis of meaning and reasoning in algebra above the too common stress on manipulative skill in complicated and unusual algebraic forms; (7) sufficient economy of time, through simplification in algebra, to introduce in the ninth year logarithms, the elements of trigonometry, and some formal demonstrative geometry; (8) constant aims at greater reality through correlation with the pupils' common life experiences; (9) an organization of subject matter with greater reference to psychological and pedagogical order than to abstract logical order; (10) some attention given to historical setting.

Teachers who are sensitive to suggestions on ways to improve high-school mathematics, even if not teaching in a junior high school, will find these good books well worth examining.

ERNEST B. LYTLE

University of Illinois

GAYLEY, CHARLES M., YOUNG, C. C., AND KURTZ, B. P. *English poetry: its principles and progress*. New York: Macmillan Company, 1920. 728 pp.

This is a revised and enlarged edition of the book bearing the same title which first appeared in 1904. The editors have here added about one hundred and fifty

short poems from later writers, especially from those whose work mirrors the life of the twentieth century both in peace and at war. In their attempts to adapt the book more thoroughly to the needs of the high-school student, the editors have omitted the former introduction, which they called "The Principles of Poetry," and have substituted about fifty pages which they have styled "An Introduction to the Study of Poetry." This new material, which presents in lucid form what every high-school student should know of that subject, has been written by Professor Gayley in his usual attractive and stimulating manner. Both the student and the general reader will enjoy both the introduction and the good poetry that follows.

H. G. PAUL

University of Illinois

SMITH, WAYNE P. AND JEWETT, EDMUND G. *An introduction to the study of science.* New York: Macmillan Company, 1919. 509 pp.

With but few exceptions, this book contains an excellent choice of material. Commendable features are the emphasis upon experimentation and utility. The style and general presentation are not well adapted to high-school freshmen. They are much better suited to seniors. The review questions at the close of the various chapters are too largely of the memoriter type and refer back too directly to the text. Had many of them been omitted and more suggestive questions been included, the book would have been materially improved. The text is recommended for reference rather than for classroom purposes.

LEWIS W. WILLIAMS

University of Illinois

TRAFTON, GILBERT H. *Science of home and community.* New York: Macmillan Company, 1920. 557 pp.

What constitutes the proper field of general science? The author answers this question as follows: "Whether any particular science or any particular topic should form a part of the course depends on the number and value of the elements that it possesses in common with the child's life." His conclusion that from this standpoint physics and hygiene should dominate the text is not supported by any evidence. It is doubtful whether the study of such topics as the wireless telegraph and the airplane can be justified by means of the above criterion.

In general, the author has succeeded in accomplishing the aim he has set for himself. The idea of grouping material around the science of the home and the science of the community seems especially fitting for his purpose. The projects outlined for pupils and the questions suggested for class discussion are well chosen. Teachers of general science when choosing a text should give this work thorough consideration.

LEWIS W. WILLIAMS

University of Illinois

DEWEY, EVELYN, CHILD, EMILY, AND BEARDSLEY, RUML. *Methods and results of testing school children.* New York: E. P. Dutton and Company, 1920. 170 pp.

This volume is a report of the use of various psychological tests used in the measurement of certain groups of children in New York City. The children were all of

Jewish parentage and were selected in such a way that representative groups were obtained for each of the ages eight to thirteen inclusive. The tests were given to fifty boys and fifty girls of each age.

After some preliminary testing the following group tests were selected.

Picture Completion	Memory for Objects
Healy A	Healy B
Identification of Forms	Memory for Digits
Opposites	Card Sorting
Cancellation	Substitution
Cart Construction	Narrative Pictures
Rossolimo	Instruction Box
Ball and Field	Steadiness
Nail Driving	Learning Box
Needle Threading	Strength
Problem Box	Binet-Yerkes
Knox Cubes	

Exceptional care was exercised in the giving of the tests. The results, therefore, have an unusual significance. A feature of the report is the form in which the test norms are given. For each age the mean and standard deviation, together with their respective probable errors, are given. For most of the tests the norm is also given in the form of the regression equation on age. By means of this equation the standards for any age within the limitation covered by the survey can be immediately obtained. The use of the regression equation assumes a rectilinear relation between the test scores and age. No correlation tables are given by means of which the validity of this assumption can be judged accurately. In some cases the relation of the means to the ages is such as to suggest that the relation is not rectilinear.

The authors attempted to select from the tests used certain ones that would form what they call a "maturity scale." Such a scale does not differ in function from the Binet scale or group intelligence tests. The criterion of selection of individual tests for this scale is chronological age. The tests were chosen which correlated most highly with chronological age. By means of this criterion seven tests were selected to form a "maturity scale" for boys and six tests were selected to form a "maturity scale" for girls. For each of these groups of tests the regression equation is given. A child's score on the maturity scale can be found by substituting his scores on the individual tests in this equation.

In addition to the report on psychological tests brief mention is made of a social and physical study of the same pupils.

The book does not appear to be of major importance. The average school man will find little in it that is helpful. The greater portion is devoted to the reproduction of the tests and to tables and charts. The sample statistical treatment of the data collected may well serve as a model for other reports. W. S. M.

SMITH, ALICE M. *Short plays by representative authors.* New York: Macmillan Company, 1920. 318 pp.

Twelve plays are here presented, mostly of the one-act variety, selected apparently on the plan of introducing the works of many nationalities. The majority of them are here exploited for the first time in a collection of this kind. Nations represented are

Iceland, Wales, England, India, The United States, Germany, Ireland, Russia, and negro America. The twelve plays run the gamut of type, from the bald realism of the Icelandic play, *The Hraun Farm*, to Stuart Walker's inimitable fancy, *Six Who Pass While the Lentils Boil*. Sentiment is provided in Jeannette Marks' *The Merry, Merry Cuckoo*, in Tagore's *The Post Office*, and in *The Shadowed Star* by Mary Macmillan. Mastery of dramatic technic and superlative dramatic movement are represented by John Masefield's adaptation of an old saga in *The Locked Chest* and in Ludwig Fulda's *By Ourselves*. Pictures of *genre* types are not absent; the most notable are Lady Gregory's *Spreading the News* and Ridgely Torrence's strong picture of the American negro in *The Rider of Dreams*.

In addition to the catholicity of taste displayed in choosing and the variety of types given, all of the plays in the collection are actable, making the book distinctly useful to those who are interested in producing good examples of modern short plays. Since most of the plays will in all probability be new to those who encounter this book, it will prove valuable as an addition to the growing literature of the short play.

C. H. WOOLBERT

University of Illinois

News Items and Communications

This department will contain news items regarding research workers and their activities. It will also serve as a clearing house for more formal communications on similar topics, preferably of not more than five hundred words. These communications will be printed over the signatures of the authors. Address all correspondence concerning this department to Walter S. Monroe, University of Illinois, Urbana, Illinois.

New Edition of "The Hurdles"

Readers of the JOURNAL OF EDUCATIONAL RESEARCH will remember the article by Director Murray A. Dalman of the Indianapolis public schools, in the January 1920 number. In this article Mr. Dalman showed a unique use which he had made of test material. This material consisted of a series of tests in each of the principal topics in algebra. For each topic there was an exercise appropriate to each of four levels of ability among students. The editors of the JOURNAL OF EDUCATIONAL RESEARCH have frequently been asked how this test material could be obtained. Mr. Dalman writes recently that he is making up a new edition of "The Hurdles." Doubtless those who are interested may obtain copies of them from him at a nominal charge.

Reading and Spelling at Beaumont, Texas

We have lately received a copy of the report prepared by Miss Clara Mallory on tests in silent reading (Monroe) and spelling. This material shows not only the achievement in these subjects in 1920, but also a comparison with achievements in 1919. In both years the testing was done in the fall. As usually happens where a continuing policy is in operation, marked improvement is shown. For example, in spelling Miss Mallory notes, "Last year it was not possible to get results from the

third grades, as they were not able to write the words; but this year the third grades in some places are up to standard and some of them are above standard." In connection with the graphs a few paragraphs of interpretation are included in the report and some suggestions to teachers.

Conference at New Orleans

On December 16, 17, and 18, the Conference of State and Parish School Officials was held at New Orleans under the auspices of the State Department of Public Instruction. Considerable attention was given to questions of educational research. Indeed, there was a decided disposition on the part of the makers of the program to give the entire discussion a fact basis. Dr. Walter S. Monroe was the principal speaker from outside the state. He addressed the conference on the following topics: (1) measurement of general intelligence and achievement of school children; (2) standards for achievement tests and their use; (3) objectives in arithmetic; (4) equal educational opportunities; (5) the teaching of arithmetic; (6) royal roads to learning.

New Tests

The Bureau of Educational Measurements, Kansas State Normal School, Emporia, Kansas, announces for distribution the following new tests: Lunceford Number Test, Kansas Latin Test and Beach Music Test.

The following statement has been received concerning the Commonwealth Fund.

Educational Research by the Commonwealth Fund

The directors of the Commonwealth Fund having become convinced of the importance of encouraging educational research requested a group of leading educational men of the country to report upon the opportunities in the field and recommend a plan of operation.

This Educational Research Conference was held in Atlantic City on October 23, 24, and 25, 1920. The members attending the conference were:
A. Ross Hill, President, University of Missouri, *Chairman*
Lotus D. Coffman, President, University of Minnesota
Charles H. Judd, Director, School of Education, University of Chicago
Paul Monroe, Director, School of Education, Teachers College, Columbia University
Leonard P. Ayres, Director, Division of Education, Russell Sage Foundation
Samuel P. Capen, Secretary, American Council on Education
Paul H. Hanus, Professor of Education, Harvard University
Max Farrand, General Director of the Commonwealth Fund

A gratifying feature of the conference was that the members were able to agree so readily, not merely upon the large divisions of the field, which were fairly obvious, but also in selecting some of the most desirable problems to be investigated within each division.

The conference therefore made its report, choosing for its main subjects: the problem of school revenues; the evaluation, analysis, measures and standards of accomplishment of school subjects; the field of supervision; and the reorganization of public school systems. In each one of those larger divisions the conference chose a particular subject by way of illustration, suggesting how and by whom the study of that particular subject might be advisedly undertaken.

For the carrying on of this research, the conference recommended the appointment of a committee to consider and recommend projects for research, and to assume executive responsibility for supervising the carrying on of such researches as might be adopted, and it was deemed advisable that this committee might also recommend researches to be undertaken by individuals or associations.

The directors of the Commonwealth Fund accepted this report, and appointed the following committee:

Leonard P. Ayres, Vice-President, Cleveland Trust Company
Samuel P. Capen, Secretary, American Council on Education, *Secretary*
Lotus D. Coffman, President, University of Minnesota
Ellwood P. Cubberley, Professor, Stanford University
Charles H. Judd, Director, School of Education, University of Chicago
Paul Monroe, Director, School of Education, Teachers College, Columbia University
Frank E. Spaulding, Professor, Yale University
Max Farrand, General Director, The Commonwealth Fund, *Chairman ex officio*

A Study of Supervision

Superintendent M. E. Moore of Beaumont, Texas, has recently attempted to answer the following questions by addressing a questionnaire to the superintendent of schools in the class of cities indicated.

"1. How many visits should a superintendent in a city of from 30,000 to 100,000 population make to each building in the course of the year, without necessarily visiting each room?

"2. How many visits a year should he make to each schoolroom?

"3. How many supervisors should the above named cities possess?

"4. What supervisors?"

Replies were received from forty cities, thirty-three ranging in population from 30,000 to 100,000 and seven ranging from 100,001 to 325,000.

Twenty of the answers to the first question were of the form "as often as possible," "no fixed schedule," "many times," "no rule." The median practice of the remaining twenty cities is from six to nine visits per year to each school building.

In response to the second question definite replies were received from twenty-four cities. The median practice in this case is two or three visits to each schoolroom.

Concerning the information received in response to the third question, Mr. Moore says: "A glance at the distribution of supervisors shows that there is no correlation between the *size* of a city and the *number of supervisors* which it may have. For example, Schenectady with 89,000 people has more supervisors than Kansas City, Missouri with 324,000 people; or St. Joseph with 78,000 people has more supervisors than Dallas with 160,000 people; or Colorado Springs with 30,000 people has twice as many as Binghampton with 67,000 people."

"It is quite evident from the data that a city reaches what is considered a full corps of supervisors by the time it attains a population of from 30,000 to 40,000, and that its increase in supervisors thereafter will in no sense keep pace with its increase in population."

Replies were received from thirty-three of the forty cities in regard to the number of supervisors. A summary is given in the following table:

Number of Supervisors	Number of Cities
32 to 35	1
28 to 31	0
24 to 27	0
20 to 23	0
16 to 19	3
12 to 15	10
8 to 11 *Median*	10
4 to 7	9
0 to 3	0
	33

Since there are 33 cities or cases, the median is the seventeenth case, which case lies in the group of cities having from 8 to 11 supervisors; hence we conclude that a range from 8 to 11 is the proper number of supervisors for a city of from 30,000 to 100,000 population.

The answers to the fourth question which were received from the 33 cities ranging in population from 30,000 to 100,000 are summarized in the following table:

	Number of Supervisors
Health (doctors and nurses)	87.5
Attendance	55
Music	50.5
Drawing	42.5
Assistant superintendent, supervisors of intermediate grades, and supervisors of tests and measurements	37.5
Physical education	36
Sewing	29
Writing	24.5
Primary grades	20.5

Commission on Coordination of Research Agencies

The National Education Association announces the appointment of a commission whose function is "to coordinate all agencies which are studying the problems of education and to work out a policy whereby these agencies will contribute to the program advocated by the Association." This commission consists of the following:

Jesse H. Newlon, Superintendent of Schools, Denver, Colorado, *Chairman*
W. C. Bagley, Editor-in-Chief of the *Journal*, Pleasantville, N. Y.
hn W. Withers, Dean of School of Pedagogy, University of New York, New York City

J. W. Studebaker, Superintendent of Schools, Des Moines, Iowa
Mary D. Bradford, Superintendent of Schools, Kenosha, Wisconsin
Walter A. Jessup, President, State University of Iowa, Iowa City, Iowa
L. D. Coffman, President, University of Minnesota, Minneapolis, Minnesota
Edwin L. Rouse, State Normal Schools, Peru, Nebraska
J. G. Crabbe, President, State Teachers' College, Greeley, Colorado
W. W. Theisen, Director of Department of Reference and Research, Cleveland, Ohio
George Melcher, Director Bureau of Research and Efficiency, Kansas City, Missouri
Virgil E. Dickson, Director of Research and Guidance, Oakland, California
Burdette R. Buckingham, Director of Bureau of Educational Research, University of Illinois, Urbana, Illinois
Thomas H. Briggs, Professor of Secondary Education, Columbia University, New York City
Charles E. Rugh, Professor of Education, University of California, Berkeley, California
W. F. Russell, Dean of College of Education, State University of Iowa, Iowa City, Iowa
Annie Webb Blanton, State Superintendent of Public Instruction, Austin, Texas
Jessie E. Colburn, Public School No. 61, New York City
Phoebe E. Matthews, 153 Morrison Avenue, Somerville, Massachusetts
Robinson G. Jones, Superintendent of Schools, Cleveland, Ohio
C. P. Cary, State Superintendent of Schools, Madison, Wisconsin
James W. Gowans, Superintendent of Schools, Winfield, Kansas
Stuart A. Courtis, 246 Eliot Street, Detroit, Michigan
Franklin Bobbitt, Professor of Education, University of Chicago, Chicago, Illinois
Frank G. Pickell, Assistant Superintendent of Schools, Cleveland, Ohio
Albert S. Cook, State Superintendent of Schools, Baltimore Maryland

Educational Research in China

The letter, a portion of which is quoted below, has so much interested us that we think our readers will also like to read it. With reference to this letter, a friend of ours writes, "This aptly illustrates the song sung some years ago by the Mandarin Yen How in that charming musical comedy *San Toy*

'They do it in the West
So of course it must be best
And I mean to introduce it into China.'"

Mr. Francis P. Jones, who writes the letter, is principal of the Hinghwa Normal School at Hinghwa, Fukien, China. He is also a member of the Fukien Christian Educational Association. At the point where we begin our quotation from Mr. Jones' letter he has just been speaking of this association. Continuing he says:

"Before our meeting at Kuliang this summer, several missionaries had already begun the use of various tests. Mr. Irving Lacy of Yenping had translated Starch's reasoning test in arithmetic and had results from several hundred children, all of them in higher primary schools, i. e., grades five to eight. I had also given the Cleveland Survey Arithmetic Tests to all children in the six principal higher primary schools of our Hinghwa Mission. This amounted to about three hundred students. I may later be able to send you a graph or a detailed statement showing the results from this test in our schools here. As far as general arithmetic ability is concerned it shows our children to be approximately one year behind the children of the same grade in

American schools. This difference in attainment is due largely to the difficulty in getting good arithmetic teachers in China. The old Chinese schools did not consider arithmetic a proper part of the curriculum which was composed entirely of classical literature.

"We are doing our best to improve the quality of the teaching so that our children may come up to the grade of schools in America. The results which Mr. Lacy secured in his reasoning tests seem to show equal attainment between the children in our schools and those in American schools. It is, however, more difficult to draw a parallel there because the translation of problems from one language into another may result in increasing or decreasing the intrinsic difficulty of the problem.

"At the meeting of the Fukien Christian Educational Association at Kuliang this summer, the above work was reported on and the Bureau of Educational Research was established.

"Mr. F. B. Beach of Foochow has undertaken to adapt group intelligence tests for use in Chinese schools. Mr. Irving Lacy proposes to proceed with his reasoning tests. Mr. Walter Lacy of Foochow is working on geography tests and measurements and I have been asked to work on the problem of English tests for higher primary and middle or high school.

"One of the Professors in Canton Christian College, H. B. Graybill, has worked out a good dictation test in English, which tests understanding of English and ability to write it correctly. His test is composed of a series of nineteen sentences, beginning with a simple sentence like—'I see a boy,' and gradually increasing in difficulty. You will find a complete account of this test in one of the numbers of *Educational Review for China* (October, 1919).

"This is the only definite information that I can give you. Some of the higher government schools are working at this problem. The Government Teachers College at Nanking, of which Dr. P. W. Quo is President, is I think working along several lines. I have heard that tests in different subjects are also being worked out at other educational centers such as Pekin University and West China University, but I am not able to give you any definite information about them."

Section Q, A. A. A. S.

Section Q of the American Association for the Advancement of Science was in session in Chicago during December 28, 29, and 30, 1920. The meeting on December 28 which was a joint session with Section I (Psychology), was devoted to the nature of mental abilities and their measurement. The particular topics are indicated in the following program:

The Correlation of Visual and Color Memory with Art Ability. Elmer E. Jones, Northwestern University, Evanston, Illinois.

The Measurement of Silent Reading. May Ayres Burgess, Russell Sage Foundation, New York City (not present).

Observations of Eye Movements in Reading. G. T. Buswell, University of Chicago, Chicago, Illinois.

Character and Conditions of Motor Rhythm as Related to Handwriting Movement. Paul V. West. University of Wisconsin, Madison, Wisconsin.

Measurements of Spelling in the Virginia Survey. F. S. Breed, University of Chicago, Chicago, Illinois.

Corrective Work in Reading. Wm. S. Gray, University of Chicago, Chicago, Illinois.

Analyzing Industrial Educational Requirements. E. K. Strong, Jr., Carnegie Institute of Technology, Pittsburgh, Pennsylvania.

The Interpretation and Application of the Intelligence Quotient. F. N. Freeman, University of Chicago, Chicago, Illinois.

Report on College Admission Tests. L. L. Thurstone, Carnegie Institute of Technology, Pittsburgh, Pennsylvania.

The Essential Criteria of Mental and Educational Tests. M. E. Haggerty, University of Minnesota, Minneapolis, Minnesota.

Combined Mental and Educational Tests. Rudolf Pintner, Ohio State University, Columbus, Ohio.

A Group Intelligence Scale for Primary Grades. F. A. Kingsbury, University of Chicago, Chicago, Illinois.

The Influence of Heredity, Maturity, and Training on Scores in Standard Tests. S. A. Courtis, Detroit Public Schools, Detroit, Michigan.

One's estimate of any program depends largely upon his personal interests. Thus in attempting to note the high points one doubtless reflects as much of his own personal interests as of the inherent nature of the numbers on the program. The following reports, however, appeared to stand out as distinctive.

Dean Gray reported in detail a case study of a fourth-grade pupil who was found to be extremely deficient in reading. His report together with other reports indicates that practically all failures in reading, except when due to a lack of native intelligence, can be eliminated by providing proper instruction for the pupils. The correction of reading defects, however, requires a careful study of the needs of the individual pupils and the provision of appropriate remedial instruction.

Professor Freeman presented data to show that there was considerable doubt whether the assumption that the I. Q. is constant is justified when it is based upon mental age as measured by group intelligence tests.

Dean Haggerty presented four criteria for judging mental and educational tests: (1) discrimination—a satisfactory test must show differences in ability where real differences in ability exist; (2) reliability—a second application of a test should yield measures which closely approximate the measures obtained from the first application; (3) significance—the traits measured should be significant, i. e., they should be included in the educational objectives; (4) standards—tests should be standardized both with respect to age (chronological) and school grade. Our present grade standards are inadequate. In the discussion which followed Dr. Thurstone suggested that the first three of these criteria might be included under the head of "validity" and that this characteristic of a test could be studied directly by correlating the measures yielded by it with an independent criterion of the trait covered by the test. This comment doubtless grew out of Dr. Thurstone's work with tests for admission to college where an independent objective measure of a student's success is obtainable. In this case it is possible to study the validity of such tests by directly correlating the measures yielded by it with the student's future success in college. This procedure cannot be applied when independent objective measures are not available. This is the situation with which we have to deal in the case of most tests of intelligence and school achievement.

Professor Pintner reported a plan for combining measures of intelligence and achievement in order to interpret more accurately the measures of achievement.

The keynote of this day's session as well as of the other sessions was voiced by the chairman, Professor Chas. H. Judd, when he pointed out that it was clearly evident that we should proceed very cautiously in interpreting the measures yielded by our present educational tests and that there is a distinct need for careful and detailed analysis before we can arrive at final conclusions. The uncritical use of tests is very likely to lead to erroneous conclusions.

On December 29 a joint session was held with the American Psychological Association. Since it is expected that a separate report of this meeting will be published in these columns it will be omitted from this account.

The program of Thursday morning, December 30, consisted of five papers. The first, by Professor B. R. Buckingham, was on a "Brief Method of Evaluating Test Material." Recognizing the need of a greater volume of such material and the difficulty of evaluating it by the direct method, the speaker considered the possibility of utilizing data already in hand to standardize educational material. It was an attempt to short-circuit the usual procedure and to investigate the results of an indirect method. He found that by combining twenty-five spelling words of undetermined difficulty with twenty-five whose difficulty was known he could determine the difficulty of the new words by using returns from three hundred children with an error amounting in no case to more than two columns on the Ayres scale, and in most cases to less than one column.

Dr. Carter Alexander, Assistant State Superintendent, State Department of Public Instruction of Wisconsin, set forth in a paper "Motives Successfully Utilized in Publicity" some of the social and psychological principles which ought to be observed in such appeals to the people. Dr. Alexander particularly depricated the appeal to *pity*. He gave instances in which emotional cartoons and stories depicting the sad condition of the teacher entirely failed to produce the desired effect.

Although Dr. W. Randolph Burgess' paper was entitled "Teacher Preparation in Ten States," he had secured comparable data for fourteen states. The percent of college and normal-school graduates among the teaching forces in these states was taken as the index for teacher preparation. The states in order in regard to the proportion of trained teachers were: Massachusetts, Rhode Island, New Jersey, Connecticut, New Hampshire, Colorado, Minnesota, Montana, Wisconsin, Illinois, Virginia, North Dakota, West Virginia, and Kansas. The percent of teachers who were not graduates of either college or normal schools ranged from fifteen for Massachusetts to seventy-four for Kansas. The eastern states showed high records and the southern and middle states lower records. A more fundamental condition, however, was that the training of teachers was most extensive in states where urban conditions most largely prevail. An index number derived for each state on the basis of the number of trained teachers revealed a correlation between teacher preparation and the percent of urban population amounting to $+0.89$. Indeed, as Dr. Burgess pointed out, the problem of teacher training is largely a rural problem. The typical city teacher is a trained teacher. About half of the total number of teachers of these fourteen states were normal-school or college graduates. Among the ten states for which data were available over a ten-year period, there was an increase of 40 percent in the amount of teacher preparation between 1910 and 1920—although in the years 1919 and 1920 little progress was made. If present rates of progress are continued until 1950, the states now having the best record will have an average of nearly four years of college

or normal school training per teacher. The present average for the best states—Massachusetts, Rhode Island, and New Jersey—is two years.

Professor A. Caswell Ellis of the University of Texas presented a paper on "The Rating of University Professors." Professor Ellis had devised by an amalgamation of the judgment of college and university professors and students a score card for judging college teachers.

Owing to unavoidable delays, the paper of Professor Fletcher Harper Swift of the University of Minnesota which was scheduled for the morning session was postponed until afternoon. Professor Swift illustrated his paper by the use of a number of especially interesting slides. He set forth in considerable detail the development of school expenditures (with emphasis upon the sharp upward trend of the curve in recent years), the various sources of school revenue, the productiveness of these sources, and the relation between local, county, state, and federal sources. Professor Swift expressed the conviction that in the present state of our educational development the only agency which can adequately finance the schools and make them in fact, as they are in theory, universal, democratic, and free, is the state government.

On the afternoon of December 30 Professor L. M. Terman of Leland Stanford Junior University read a report by Dr. Grace Fernald on the teaching of children to read who were afflicted with aphasia or word blindness. Several cases were reported all of which were treated successfully. A series of articles on this topic by Dr. Fernald will appear soon in the JOURNAL OF EDUCATIONAL RESEARCH.

Professor Bird T. Baldwin, State University of Iowa, presented data which showed that the anatomical and psychological ages of children should receive consideration as well as chronological and mental ages in classifying them for instruction. A complete account of Professor Baldwin's work will appear soon in book form.

Mr. L. J. Brueckner of the Detroit public schools reported on their method of teaching English to adult foreigners. The essentials of the method involved using material which had practical content.

Mrs. Luella Pressey of Indiana University reported a critical study of Monroe's Standardized Silent Reading Tests. She found that when a test of this type was restricted to exercises based on prose paragraphs the correlation of one test with a similar test was very high (0.85). For the same group of pupils the correlation with similar tests based on poetry and upon selections from general science was 0.35 and 0.39 respectively. The conclusion suggests that different abilities are required for reading different types of material and that the validity of a test of this type can be greatly increased by confining the exercises to a single type of reading material.

Spokane U. S. History Test

The Spokane U. S. History test was devised in the latter part of the year 1919, with a view to correcting certain defects in other standardized history tests. It is so arranged that guessing is practically eliminated. It is rather searching, having one hundred points on which answers are required, and yet requires an hour or less to give. The scoring of the papers is easy and exact, leaving nothing for the teacher to guess at. It is fairly complete as it is made up of twenty points in each of five sections, dealing respectively with dates, people, historic expressions, civic terms and causes in their relation with historic events.

In the first section twenty important dates are placed along the left-hand margin of the page and the twenty historic events with which they are associated are placed

in a different order down the center of the page. On the right of the page is a vacant column headed "Date." The instructions to pupils are, "Select dates from the column on the left and write them in the vacant column on the right, placing each date opposite the historic event associated with it."

A similar arrangement holds true of each of the other sections. The second section is here reproduced entire in order to illustrate the plan.

II

Select names from the column on the left and write them in the vacant column on the right, placing each name opposite the historic event associated with it.

Name	Historic Event	Name
Peary	Destroyed Spanish fleet in Manila Bay	
Sacajawea	Arranged slavery compromises	
Foch	Invented the telephone	
Lincoln	Leading Confederate General	
Fulton	Wrote the Declaration of Independence	
Dewey	Built the Panama Canal	
Pershing	Discovered North American Continent	
Roosevelt	Wrote the "Star Spangled Banner"	
Bell	Invented the steamboat	
Washington	Guided the Lewis and Clark Expedition	
Goethals	Discovered the North Pole	
Magellan	Commanded allied armies in the World War	
Edison	Was President during Civil War	
Grant	Commanded American Forces in World War	
Clay	Was Revolutionary patriot, author and inventor	
Jefferson	Commanded "Rough Riders" in Spanish American War	
Lee	Discovered the Philippine Islands	
Key	First president of the United States	
Cabot	Leading Union General in Civil War	
Franklin	America's most famous inventor	

No. Attempted..... No. Right.....

The test was given in January 1920, to the 514 pupils completing the eighth grade at the midyear. The medians for the 28 buildings in which there were VIIIA classes ranged from 59.0 to 93.5. The city median was 78. The quartile medians were 59 to 69, 70 to 78, 78 to 82.5 and 83.5 to 93.5.

The boys ranked higher than the girls in all sections. The median for 236 boys was 81 and for 298 girls was 74.5. The lowest score was 27 and was made by a girl. The highest score was 100 and was made by one boy and two girls. Thirty-three pupils, or 6.4 percent, had a mark of 50 or less. As this is the average percent of failure in the Spokane schools, it was decided to call the passing mark 51 or above.

The results by ages correspond to those nearly always found, namely, that the younger pupils in a class are the stronger ones. One pupil eleven-years-old received

a mark of 92. Twelve pupils twelve-years-old had a median mark of 90. One hundred four pupils thirteen-years-old had a median mark of 84.5. Two hundred six pupils fourteen-years-old had a median of 77. One hundred thirty-three pupils fifteen-years-old had a median of 76. Forty-nine pupils sixteen-years-old had a mark of 71. Seven pupils seventeen-years-old had a mark of 61. Two pupils eighteen-years-old had a mark of 73.

Of a possible mark of 20 for each section, the median marks were 16 for dates, 20 for people, 17 for historic terms, 16 for civic terms, and 11 for cause and effect. It appears, therefore, that associations with persons are easiest and those of cause and effect most difficult.

The five easiest points in the order of easiness were (1) the name of the agreement ending the fighting in the World War; (2) the name of the first president of the United States; (3) the date of the beginning of the World War; (4) the name of the ship on which the Pilgrims came; and (5) the name of the boundary between the slave and free states.

The most difficult points in the order of difficulty were: (1) disputes over slavery as causing the Civil War; (2) the Civil War as bringing about national unity; (3) the industrial development of the New England states as giving rise to a protective tariff; (4) the lack of a strong, centralized government as the cause of the critical period; and (5) the invention of the steam engine and machinery as the cause of the growth of cities.

In tabulating the results of the test a separate card, 3 by 5 inches in size, was used for each pupil. On this card was entered the name, age, sex and building, together with the mark of the pupil in each of the five sections, and the total. The use of these cards made possible the easy segregation of the results and finding of medians. In the opinion of the writer the use of such cards is to be recommended as a great value in any similar work.

The use of the test seems to show that it possesses the following advantages:

1. Pupils understand what to do from reading the brief instructions given in the test itself, thus making unnecessary any explanation by the teacher.

2. The key accompanying the test makes it quick and easy to score, and eliminates any uncertainty as to the mark to be given.

3. It is our conviction after using the test that it points unerringly to the strength or weakness in the history work of the individual pupil or class.

4. The test is diagnostic in that a detailed comparison of individual or class results points out specific weaknesses.

ORVILLE C. PRATT

Superintendent of Schools
Spokane, Washington

Validity of Illinois General Intelligence Scale

Superintendent L. W. Keeler, Michigan City, Indiana, has furnished the Bureau of Educational Research with data from which the coefficient of correlation between scores yielded by the Illinois General Intelligence Scale and the Otis Group Intelligence Scale has been calculated. The coefficient of correlation for 124 pupils in the lower sixth grade is 0.83. For 83 pupils in the upper sixth grade it is 0.82. The Otis Group Intelligence Scale was given in September 1920, and the Illinois scale about two months later. The Otis scale had previously been given to

several of the pupils in the upper sixth grade in May 1920. The correlation between these two sets of scores for 64 pupils was found to be 0.76.

These coefficients of correlation are based upon I. Q.'s rather than point scores. They show a high degree of correlation between the measures yielded by these two instruments for measuring intelligence. This suggests that the two scales measure the same traits of pupils.

In the lower sixth grade the average I. Q. for the point score was 90.5. For the Illinois scale it was 89.9. In the upper sixth the averages were—Otis 106.2, Illinois 94.2. These averages suggest that the I. Q.'s obtained from the use of the Illinois scale are lower than the corresponding I. Q.'s for the Otis Scale. This in turn means that the mental ages yielded by the Illinois Scale are lower. Other data gathered by the Bureau of Educational Research indicate that the mental ages obtained by the Illinois scale are about one year and two months too low. This estimate is subject to revision on the basis of additional data which are now being collected. The use of this correction will insure more accurate mental ages. However, this corrected mental age should not be used in determining achievement quotients. To do so will introduce a considerable error in the quotients.

W. S. MONROE

National Association of Directors of Educational Research

(E. J. ASHBAUGH, *Secretary and Editor*)

Carter Alexander, Assistant State Superintendent, Wisconsin, reports that his research energies have been exerted in the following channels: publicity campaigns for better school support, state aid provisions, and the formulation of legislation needed in Wisconsin.

He calls attention to articles in the June JOURNAL OF EDUCATIONAL RESEARCH; July, August, and September *American School Board Journal;* and October *Educational Review;* also, the December *School and Society*. He also announces an article to appear shortly in the *School Review* and the *Elementary School Journal*.

Mrs. Margaret S. Brainard, Director of Educational Research, Martins Ferry, Ohio, states that her work this year has been largely along the lines of remedial measures following up the conditions revealed by the standard tests given for the past two years. She has been working with her superintendent on a bulletin of methods and devices for teachers of reading in the elementary grades. They have checked over the school library and prepared graded book lists as a stimulus toward home reading. She has planned to give the Thorndike Visual Vocabulary Test to discover the part played by deficient vocabularies.

Some intelligence testing has been done with both the Haggerty Group Test and the Binet Individual. Further testing with subject-matter and intelligence tests will

be made for the purpose of classification and study of progress of pupils doubly promoted.

She also reports the formation of an Ohio Educational Research Association at a meeting held at the university in December.

Dean Chas. Fordyce, University of Nebraska, reports the following items:

1. Group mental tests were given to all students of Teachers College of the university during the first semester. Results will be used with the grades given by instructors as a basis of determining scholastic standing.

2. Have used various educational tests, Haggerty Group Tests in grades, and the Terman Group Tests in the high school in several public schools.

3. Are giving the Ayres Spelling to rural schools of four counties for the purpose of establishing a norm in their state.

4. Making a comparative study of the Henmon, Starch, and Hanus Latin tests in four different public schools.

Miss Mary Bess Henry, Director of Research, Santa Ana, California, sends the following:

"The Santa Ana schools have extended the classification of pupils into groups according to ability and application. This system has been used for three years in our sixth grade and is now used in all classes from the sixth grade through the junior high, and a few in the senior high. The work expected of the slowest sections is quite different from that of the rapid sections. All children in the fifth, sixth, seventh and eighth grades have been given two group tests, Otis, Illinois Examination, National Intelligence or Terman, as well as standard tests in reading, arithmetic, spelling and writing. Trabue has been given in many cases as a third test. These records, with his school marks, record of health, vision and hearing, and estimates made by his five or six teachers of intelligence, temperament, dependability, initiative, ambition, neatness and conduct are sent to the grade or school to which he is promoted.

"Three special rooms for atypical children are maintained, with some attempt at grading. Two rooms care for the younger children. They progress in time to the central room, where one half the time is spent in manual training, sewing, and cooking. The usual handwork is prominent in all the special rooms. From the central room, the higher grade cases progress to the slowest sections of the seventh and eighth grades, where they have a differentiated course of study. These promotions have a very advantageous effect upon the high grade morons and borderline cases in the special rooms. A promotion seems to mean a great deal more to them than we have been led to expect."

Dr. W. A. McCall, Teachers College, Columbia, tells us that he and Dr. Trabue, under the direction of Dr. Strayer, are in charge of mental measurement for the Baltimore school survey which is now in progress. Also that form 1 of the Thorndike-McCall Reading Scale is ready for distribution. The scoring and computation of pupil and class scores have been very much simplified. He announces an article on method of scale construction in the January issue of the *Teachers College Record*, and a forthcoming volume on "How to Measure in Education" to be published by Macmillan.

Dr. M. R. Trabue, of Teachers College, announces the following lines of activity. He is directing studies of the different intelligence tests as to their relative values in:

1. Classifying pupils entering junior high school.
2. Determining the character of pupils eliminated from metropolitan high schools.
3. Placement of their own graduates.

He states he has a group of superintendents and principals from New York, New Jersey, and Connecticut who come in for criticism and instructions regarding tests and measurement work on Saturdays.

He also states that he and Dr. McCall have used the following pedagogical tests, which were not mentioned by Dr. McCall, in the Baltimore Survey: Courtis Arithmetic, Ayres Spelling, Thorndike Handwriting, Hillegas Composition Scale.

We are glad to announce the following new members to our Association during the past month:

Miss Margaret V. Cobb, Secretary of the Bureau of Mental Tests and Measurements under G. M. Whipple, University of Michigan. Miss Cobb has an A.B. from Radcliffe and an A.M. from Illinois. She has practically completed her work for the doctor's degree in the field of mental testing.

Miss Helen Davis, Director of the Bureau of Research, Jackson, Michigan. A.B. University of Illinois, with graduate work in University of Illinois and Carnegie Institute of Technology, specializing in the field of tests and measurements. Miss Davis states the following program for the work of her Bureau: Under her direction the teachers have given the National Intelligence Tests to all children from IIIB to VIA inclusive. Recommendations for classification will be made on the basis of these results. She is conducting an intensive course for kindergarten teachers in the use of the Binet-Simon test. Teachers will give this examination to all their pupils before the next semester, and all children entering the first grade in February will be classified on the basis of mental age. She has given about one hundred Binet examinations for the purpose of reclassification of pupils. Next semester it is planned to follow up work in the lower grades and testing in intermediate and high schools. Test results in grades above the sixth will be used not only for classification but for educational and vocational guidance.

Mr. Osborne Williams, Director of the Department of Educational Research and Vocational Guidance, Atlanta, Ga. A.M. Transylvania University, with some graduate work in the University of Chicago. He was a member of the Psychological Examining Board, Camp Hancock, for six months and assistant in Psychological Service, Walter Reed Hospital, for six months. At present the department is giving the Otis and the National Intelligence tests to the high-school freshmen, and some arithmetic and reading tests in the upper grades of the grammar school.